

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2005, 09:52:06 ; Search time 156 Seconds  
(without alignments)

50.107 Million cell updates/sec

Title: US-10-089-273-7

Perfect score: 107

Sequence: 1 TPEKERTAAPAEPAPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1755696 seqs, 390834859 residues

Total number of hits satisfying chosen parameters: 1755696

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 107   | 100.0       | 150    | 10 | US-09-847-208-106    |
| 2          | 68    | 63.6        | 36946  | 18 | US-10-840-512-155    |
| 3          | 62    | 57.9        | 137    | 16 | US-10-425-115-278305 |
| 4          | 61    | 57.0        | 329    | 16 | US-10-437-963-168229 |
| 5          | 61    | 57.0        | 666    | 9  | US-09-791-171-70     |
| 6          | 61    | 57.0        | 666    | 10 | US-09-804-980-70     |
| 7          | 61    | 57.0        | 666    | 14 | US-10-080-170-640    |
| 8          | 61    | 57.0        | 666    | 15 | US-10-282-122A-64944 |
| 9          | 61    | 57.0        | 666    | 16 | US-10-620-246-70     |
| 10         | 61    | 57.0        | 666    | 16 | US-10-080-170-640    |
| 11         | 61    | 57.0        | 666    | 16 | US-10-468-356-640    |

|    |      |      |      |    |                      |                   |
|----|------|------|------|----|----------------------|-------------------|
| 12 | 60.5 | 56.5 | 136  | 16 | US-10-437-963-104356 | Sequence 104356,  |
| 13 | 60.5 | 56.5 | 672  | 16 | US-10-425-115-231536 | Sequence 231536,  |
| 14 | 59.5 | 55.6 | 1585 | 13 | US-10-087-192-651    | Sequence 651, App |
| 15 | 59.5 | 55.6 | 2462 | 10 | US-09-819-104A-5     | Sequence 5, Appli |
| 16 | 59   | 55.1 | 20   | 11 | US-09-739-466C-24    | Sequence 24, Appl |
| 17 | 59   | 55.1 | 221  | 14 | US-10-158-761-12019  | Sequence 12019, A |
| 18 | 59   | 55.1 | 382  | 14 | US-10-152-886-31     | Sequence 31, Appl |
| 19 | 59   | 55.1 | 382  | 20 | US-11-053-576-31     | Sequence 31, Appl |
| 20 | 59   | 55.1 | 382  | 20 | US-11-053-052-31     | Sequence 62042, A |
| 21 | 59   | 55.1 | 428  | 15 | US-10-282-122A-62042 | Sequence 206940,  |
| 22 | 58   | 54.2 | 188  | 16 | US-10-425-115-253209 | Sequence 253209,  |
| 23 | 58   | 54.2 | 200  | 16 | US-10-425-115-332007 | Sequence 332007,  |
| 24 | 57.5 | 53.7 | 103  | 16 | US-10-425-115-332007 | Sequence 40572, A |
| 25 | 57.5 | 53.7 | 381  | 16 | US-10-767-701-40572  | Sequence 74781, A |
| 26 | 57   | 53.3 | 379  | 15 | US-10-282-122A-74781 | Sequence 138, App |
| 27 | 57   | 53.3 | 379  | 16 | US-10-474-792-138    | Sequence 107, App |
| 28 | 57   | 53.3 | 497  | 14 | US-10-339-217-107    | Sequence 2, Appli |
| 29 | 57   | 53.3 | 764  | 14 | US-10-399-140-2      | Sequence 231793,  |
| 30 | 57   | 53.3 | 894  | 16 | US-10-425-115-231793 | Sequence 7036, Ap |
| 31 | 57   | 53.3 | 949  | 17 | US-10-732-923-7036   | Sequence 2, Appli |
| 32 | 57   | 53.3 | 972  | 16 | US-10-484-148-2      | Sequence 4, Appli |
| 33 | 57   | 53.3 | 996  | 15 | US-10-380-492A-4     | Sequence 2, Appli |
| 34 | 57   | 53.3 | 1002 | 15 | US-10-380-492A-2     | Sequence 2, Appli |
| 35 | 57   | 53.3 | 1009 | 14 | US-10-290-544-2      | Sequence 5, Appli |
| 36 | 57   | 53.3 | 1009 | 14 | US-10-290-544-5      | Sequence 3123, Ap |
| 37 | 57   | 53.3 | 1023 | 15 | US-10-108-260A-3123  | Sequence 8, Appli |
| 38 | 57   | 53.3 | 1023 | 15 | US-10-380-492A-8     | Sequence 143815,  |
| 39 | 56.5 | 52.8 | 158  | 15 | US-10-424-599-143815 | Sequence 294429,  |
| 40 | 56.5 | 52.8 | 165  | 16 | US-10-425-115-294429 | Sequence 6369, Ap |
| 41 | 56.5 | 52.8 | 452  | 15 | US-10-369-493-6369   | Sequence 299493,  |
| 42 | 56.5 | 52.8 | 1460 | 16 | US-10-425-115-259987 | Sequence 58987, A |
| 43 | 56   | 52.3 | 132  | 16 | US-10-767-701-58987  | Sequence 42062, A |
| 44 | 56   | 52.3 | 149  | 16 | US-10-767-701-42062  | Sequence 57, Appl |
| 45 | 56   | 52.3 | 181  | 15 | US-10-299-636-57     |                   |

#### ALIGNMENTS

RESULT 1  
US-09-847-208-106  
; Sequence 106, Application US/09847208  
; Publication No. US20030082190A1  
; GENERAL INFORMATION:  
; APPLICANT: Saxon, Andrew  
; APPLICANT: Zhang, Ke  
; APPLICANT: Zhu, Daoheng  
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES  
; FILE REFERENCE: UC67.002A  
; CURRENT APPLICATION NUMBER: US/09/847,208  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 106  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Hevea brasiliensis (Para rubber tree)  
US-09-847-208-106

Query Match 100.0%; Score 107; DB 10; Length 150;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPEKERTAAPAEPAPE 20

Db 45 TPEKERTAAPAEPAPE 64

#### RESULT 2

US-10-840-512-155  
; Sequence 155, Application US/10840512  
; Publication No. US20050125852A1

```
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 155
; LENGTH: 36946
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-840-512-155

Query Match      53.6%; Score 68; DB 18; Length 36946;
Best Local Similarity 68.4%; Pred. No. 9.6e+02;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 PEKEEPTAAPAEPEAPAE 20
Db      13709 PEKLPVAAAPKPEAPAE 13727

RESULT 3
US-10-425-115-278305
; Sequence 278305, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 278305
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(137)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_185389C.1.pap
US-10-425-115-278305

Query Match      57.9%; Score 62; DB 16; Length 137;
Best Local Similarity 57.9%; Pred. No. 17;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      1 TPEKEEPTAAPAEPEAPAE 19
Db      9 TPRAESPAASPPRPAPAP 27

RESULT 4
US-10-437-963-168229
; Sequence 168229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168229
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66764C.1.pap
US-10-437-963-168229

Query Match      57.0%; Score 61; DB 16; Length 329;
Best Local Similarity 63.2%; Pred. No. 53;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 TPEKEEPTAAPAEPEAPAE 19
Db      195 TPTATTAAAPAPAPAP 213

RESULT 5
US-09-791-171-70
; Sequence 70, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDSINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-70

Query Match      57.0%; Score 61; DB 9; Length 666;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 PEKEEPTAAPAEPEAPAE 19
Db      156 PTPQTFTGAPOQPEAPAP 173

RESULT 6
US-09-804-980-70
```

```

Query Match          57.0%; Score 61; DB 15; Length 666;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0

Qy      2 PEKEPTTAPEPEAPAP 19
      | : ||| :|||
Db      156 PTPQPTGAPOQPESAP 173

RESULT 9
US-10-620-246-70
; Sequence 70, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1A
; CURRENT APPLICATION NUMBER: US/10/620,246
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171

```

```
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-70

Query Match
Best Local Similarity 57.0%; Score 61; DB 16; Length 666;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEAPAP 19
Db 156 PTPQTGTGAPQQPESAP 173

RESULT 10
US-10-080-170-640
; Sequence 640, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 640
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-640

Query Match
Best Local Similarity 57.0%; Score 61; DB 16; Length 666;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEAPAP 19
Db 156 PTPQTGTGAPQQPESAP 173

RESULT 11
US-10-468-356-640
; Sequence 640, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
```

```
; SEQ ID NO 640
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-640

Query Match
Best Local Similarity 57.0%; Score 61; DB 16; Length 666;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEAPAP 19
Db 156 PTPQTGTGAPQQPESAP 173

RESULT 12
US-10-437-963-104356
; Sequence 104356, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104356
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(136)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101699C.1.pap
US-10-437-963-104356

Query Match
Best Local Similarity 56.5%; Score 60.5; DB 16; Length 136;
Matches 14; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 2 PEKEPTAAPAEAPAP 20
Db 54 PAAEFPKAEPEAEAPAPE 73

RESULT 13
US-10-425-115-231536
; Sequence 231536, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231536
; LENGTH: 672
; TYPE: PRT
```

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_142759C.1.pap  
US-10-425-115-231536

Query Match 56.5%; Score 60.5; DB 16; Length 672;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 15; Conservative 2; Mismatches 3; Indels 5; Gaps 2;

QY 1 TPEKEPTAAP-----APEAP-APE 20  
Db 324 TPEEAPSAQPEVPSAQPEAPSAPE 348

RESULT 14  
US-10-087-192-651  
; Sequence 651, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 651  
; LENGTH: 1585  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-192-651

Query Match 55.6%; Score 59.5; DB 13; Length 1585;  
Best Local Similarity 63.2%; Pred. No. 3.6e+02;  
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPPEAPAPE 20  
Db 484 PPPEEPAAPAP-SPVPD 501

RESULT 15  
US-09-819-104A-5  
; Sequence 5, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 2462  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-819-104A-5

Query Match 55.6%; Score 59.5; DB 10; Length 2462;  
Best Local Similarity 63.2%; Pred. No. 5.6e+02;  
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPPEAPAPE 20

Db 778 PPPEEPAAPAP-SPVPD 795

RESULT 16  
US-09-739-466C-24  
; Sequence 24, Application US/09739466C  
; Publication No. US20050107585A1  
; GENERAL INFORMATION:  
; APPLICANT: MURRAY, JOSEPH S  
; APPLICANT: SIAHAAN, TERUNA J  
; APPLICANT: HU, YONGBO  
; TITLE OF INVENTION: SIGNAL-1/SIGNAL-2 BIFUNCTIONAL PEPTIDE INHIBITORS  
; FILE REFERENCE: 23902-08805  
; CURRENT APPLICATION NUMBER: US/09/739,466C  
; CURRENT FILING DATE: 2000-12-18  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 24  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Hevia brasiliensis  
US-09-739-466C-24

Query Match 55.1%; Score 59; DB 11; Length 20;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAP 11  
Db 10 TPEKEPTAAP 20

RESULT 17  
US-10-156-761-12019  
; Sequence 12019, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 12019  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-12019

Query Match 55.1%; Score 59; DB 14; Length 221;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPPEAPAPA 18  
Db 19 TPEDAEKAPAPSEGAAPA 36

RESULT 18  
US-10-152-886-31  
; Sequence 31, Application US/10152886  
; Publication No. US20030064491A1

```
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; TITLE OF INVENTION: STRUCTURES
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/10/152,886
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Streptomyces ghanaensis
US-10-152-886-31

Query Match          55.1%; Score 59; DB 14; Length 382;
Best Local Similarity 61.1%; Pred. No. 1e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 PEKEEPTAAPAEAPAP 19
DB      318 PSGAQPDAAQAQAEAPAP 335

RESULT 19
US-11-053-576-31
; Sequence 31, Application US/11053576
; Publication No. US20050142601A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; TITLE OF INVENTION: STRUCTURES
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/11/053,576
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US/10/152,886
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Streptomyces ghanaensis
US-11-053-576-31

Query Match          55.1%; Score 59; DB 20; Length 382;
Best Local Similarity 61.1%; Pred. No. 1e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 PEKEEPTAAPAEAPAP 19
DB      318 PSGAQPDAAQAQAEAPAP 335

RESULT 20
US-11-053-052-31
; Sequence 31, Application US/11053052
; Publication No. US20050170411A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; TITLE OF INVENTION: STRUCTURES
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/11/053,052
```

```
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US/10/152,886
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Streptomyces ghanaensis
US-11-053-052-31

Query Match          55.1%; Score 59; DB 20; Length 382;
Best Local Similarity 61.1%; Pred. No. 1e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 PEKEEPTAAPAEAPAP 19
DB      318 PSGAQPDAAQAQAEAPAP 335

RESULT 21
US-10-282-122A-62042
; Sequence 62042, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62042
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62042

Query Match          55.1%; Score 59; DB 15; Length 428;
Best Local Similarity 78.9%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
```



; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 74781  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-10-282-122A-74781

Query Match 53.3%; Score 57; DB 15; Length 379;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20  
Db 221 SPESEEPSVAAPSEETPSPE 240  
:|||||:|:|:|

RESULT 27  
US-10-474-792-138  
; Sequence 138, Application US/10474792  
; Publication No. US20040236072A1  
; GENERAL INFORMATION:  
; APPLICANT: Olmsted, Stephen  
; APPLICANT: Zagursky, Robert  
; APPLICANT: Nickbarg, Elliot  
; APPLICANT: Winter, Lourie  
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES  
; FILE REFERENCE: AM 100399  
; CURRENT APPLICATION NUMBER: US/10/474,792  
; CURRENT FILING DATE: 2003-10-14  
; NUMBER OF SEQ ID NOS: 674  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 138  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-10-474-792-138

Query Match 53.3%; Score 57; DB 16; Length 379;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;

Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 TPEKEPTAAPAEPEAPE 20  
Db 221 SPESEEPSVAAPSEETPSPE 240  
:|||||:|:|:|

## RESULT 28

US-10-339-217-107  
; Sequence 107, Application US/10339217  
; Publication No. US20030198648A1  
; GENERAL INFORMATION:  
; APPLICANT: Buechter, Douglas  
; APPLICANT: Hou, Xiaohong  
; APPLICANT: Marlor, Christopher W.  
; APPLICANT: Rice, William G.  
; APPLICANT: Yang, Wengang  
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of  
; FILE REFERENCE: 111021.143(ACH-US1)  
; CURRENT APPLICATION NUMBER: US/10/339,217  
; CURRENT FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: US 60/347,369  
; PRIOR FILING DATE: 2002-01-11  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 107  
; LENGTH: 497  
; TYPE: PRT  
; ORGANISM: HIV-1  
US-10-339-217-107

Query Match 53.3%; Score 57; DB 14; Length 497;  
Best Local Similarity 64.7%; Pred. No. 2.2e+02;  
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 KEEPTAAPAEPEAPE 20  
Db 447 REEPTAAPAEPTAPAE 463  
:|||||:|:|:|

RESULT 29  
US-10-399-140-2  
; Sequence 2, Application US/10399140  
; Publication No. US20030165973A1  
; GENERAL INFORMATION:  
; APPLICANT: Bayer AG  
; TITLE OF INVENTION: REGULATION OF HUMAN PGC-1-LIKE PROTEIN  
; FILE REFERENCE: L10187 Foreign Countries  
; CURRENT APPLICATION NUMBER: US/10/399,140  
; CURRENT FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: US 60/240,844  
; PRIOR FILING DATE: 2000-10-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-140-2

Query Match 53.3%; Score 57; DB 14; Length 764;  
Best Local Similarity 52.6%; Pred. No. 3.3e+02;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEPEAPE 20  
Db 119 PSSAPPSPAEKPSAPAE 137  
:|||||:|:|:|

RESULT 30  
US-10-425-115-231793  
; Sequence 231793, Application US/10425115



```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231793
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_142990C.1.pep
US-10-425-115-231793

Query Match      53.3%; Score 57; DB 16; Length 894;
Best Local Similarity 53.8%; Pred. No. 3.9e+02;
Matches 14; Conservative 3; Mismatches 3; Indels 6; Gaps 2;

Qy 1 TPEKEPTAAP-----APEAPA--PE 20
    |||||:|:|:|:|:|:|:|:|:|
Db 463 TPEPEAPSAQPEATSAQPEAPSAPPE 488

RESULT 31
US-10-732-923-7036
; Sequence 7036, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/7310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 7036
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Desulfovibrio desulfuricans G20
US-10-732-923-7036

Query Match      53.3%; Score 57; DB 17; Length 949;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EKEEPTAAPAPPEAP 17
    |||||:|:|:|:|:|:|
Db 908 ERQBPAPAPPEAP 922

RESULT 32
US-10-484-148-2
; Sequence 2, Application US/10484148
; Publication No. US20040248251A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti G.; HONCHELL, Cynthia D.;
; APPLICANT: FORSYTHE, Ian J.; CHAWLA, Narinder K.;
; APPLICANT: TANG, Y. Tom; BOROWSKY, Mark L.; BARROSO, Ines;
; APPLICANT: YUE, Henry; WARREN, Bridget A.;
; APPLICANT: THANGAVELU, Kavitha; GIETZEN, Kimberly J.;
; APPLICANT: AZIMZAI, Yalda; LEE, Ernestine A.;
; APPLICANT: BAUGHN, Mariah R.; GORVAD, Ann E.;
; APPLICANT: DUGGAN, Brendan M.; TRAN, Bao;
; APPLICANT: LI, Joana X.; RICHARDSON, Thomas W.;
; APPLICANT: ELLIOTT, Vicki S.; ZEBARJADIAN, Yeganeh
; APPLICANT: TRAN, Uyen K.; YAO, Monique G.;
```

```
; APPLICANT: PETERSON, David P.; LUO, Wen
; APPLICANT: LEHR-MASON, Patricia M.
; TITLE OF INVENTION: RECEPTORS AND MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1082 USN
; CURRENT APPLICATION NUMBER: US/10/484,148
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/22833
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US 60/306,020
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/308,179
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/309,702
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/311,476
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/311,718
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/311,551
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/314,798
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,639
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,996
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 70475510CD1
US-10-484-148-2

Query Match      53.3%; Score 57; DB 16; Length 972;
Best Local Similarity 52.6%; Pred. No. 4.2e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PEKEPTAAPAPPEAPE 20
    |||||:|:|:|:|:|:|
Db 119 PSSAPPAPAPPEAPPE 137

RESULT 33
US-10-380-492A-4
; Sequence 4, Application US/10380492A
; Publication No. US20040077536A1
; GENERAL INFORMATION:
; APPLICANT: Hart, Kevin Anthony
; APPLICANT: Montague, Carl T.
; APPLICANT: Vidal-Puig, Antonio
; TITLE OF INVENTION: HUMAN AND RAT PGC-3, PPAR-GAMMA COACTIVATIONS AND SPLICE VARIANT
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: ASZD-P01-177
; CURRENT APPLICATION NUMBER: US/10/380,492A
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: PCT/GB01/04074
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: GB 0022670.4
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 996
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-380-492A-4

Query Match      53.3%; Score 57; DB 15; Length 996;
Best Local Similarity 52.6%; Pred. No. 4.3e+02;
```

```
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 2 PEKEEPTAAPAEPAPE 20
| : ||: |||||
Db 110 PSSAPSPAPEKPSAPE 128

RESULT 34
US-10-380-492A-2
; Sequence 2, Application US/10380492A
; Publication No. US20040077536A1
; GENERAL INFORMATION:
; APPLICANT: Hart, Kevin Anthony
; APPLICANT: Montague, Carl T.
; APPLICANT: Vidal-Puig, Antonio
; TITLE OF INVENTION: HUMAN AND RAT PGC-3, PPAR-GAMMA COACTIVATIONS AND SPLICE VARIANT
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: AS2D-P01-177
; CURRENT APPLICATION NUMBER: US/10/380,492A
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: PCT/GB01/04074
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: GB 0022670.4
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-380-492A-2

Query Match 53.3%; Score 57; DB 15; Length 1002;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEPAPE 20
| : ||: |||||
Db 110 PSSAPSPAPEKPSAPE 128

RESULT 35
US-10-290-544-2
; Sequence 2, Application US/10290544
; Publication No. US20030124598A1
; GENERAL INFORMATION:
; APPLICANT: Spiegelman, Bruce M.
; APPLICANT: Lin, Jiandie
; TITLE OF INVENTION: PGC-1BETA, A NOVEL PGC-1 HOMOLOGUE AND USES THEREFOR
; FILE REFERENCE: DFN-039
; CURRENT APPLICATION NUMBER: US/10/290,544
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 60/338,126
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-290-544-2

Query Match 53.3%; Score 57; DB 14; Length 1009;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEPAPE 20
| : ||: |||||
Db 119 PSSAPSPAPEKPSAPE 137

RESULT 36
US-10-290-544-5
```

```
; Sequence 5, Application US/10290544
; Publication No. US20030124598A1
; GENERAL INFORMATION:
; APPLICANT: Spiegelman, Bruce M.
; APPLICANT: Lin, Jiandie
; TITLE OF INVENTION: PGC-1BETA, A NOVEL PGC-1 HOMOLOGUE AND USES THEREFOR
; FILE REFERENCE: DFN-039
; CURRENT APPLICATION NUMBER: US/10/290,544
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 60/338,126
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-290-544-5

Query Match 53.3%; Score 57; DB 14; Length 1009;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEPAPE 20
| : ||: |||||
Db 119 PSSAPSPAPEKPSAPE 137

RESULT 37
US-10-108-260A-3123
; Sequence 3123, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3123
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3123

Query Match 53.3%; Score 57; DB 15; Length 1023;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEPAPE 20
| : ||: |||||
Db 131 PSSAPSPAPEKPSAPE 149

RESULT 38
US-10-380-492A-8
; Sequence 8, Application US/10380492A
; Publication No. US20040077536A1
; GENERAL INFORMATION:
; APPLICANT: Hart, Kevin Anthony
; APPLICANT: Montague, Carl T.
; APPLICANT: Vidal-Puig, Antonio
; TITLE OF INVENTION: HUMAN AND RAT PGC-3, PPAR-GAMMA COACTIVATIONS AND SPLICE VARIANTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: AS2D-P01-177
; CURRENT APPLICATION NUMBER: US/10/380,492A
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: PCT/GB01/04074
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: GB 0022670.4
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 40
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-380-492A-8

Query Match      53.3%; Score 57; DB 15; Length 1023;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      2 PEKEPTAAPAEPEAPE 20
      | | | | | | | | | |
Db      131 PSSAPSPAPEKPSAPE 149

RESULT 39
US-10-424-599-143815
; Sequence 143815, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143815
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(158)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100879C.1.pep
US-10-424-599-143815

Query Match      52.8%; Score 56.5; DB 15; Length 158;
Best Local Similarity 57.1%; Pred. No. 80;
Matches 12; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY      3 EKEEPTAAP---AEPEAPE 20
      | | | | | | | | | |
Db      83 EQQPAAAPGDDAAPKPPAPE 103

RESULT 40
US-10-425-115-294429
; Sequence 294429, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 294429
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31604C.1.pep
```

```
US-10-425-115-294429

Query Match      52.8%; Score 56.5; DB 16; Length 165;
Best Local Similarity 57.1%; Pred. No. 83;
Matches 12; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY      3 EKEEPTAAP---AEPEAPE 20
      | | | | | | | | | |
Db      83 EQQPAAAPGDDAAPKPPAPE 103

Search completed: August 11, 2005, 10:02:01
Job time : 157 secs
```

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2005, 09:39:41 ; Search time 171 Seconds  
(without alignments)

59.892 Million cell updates/sec

Title: US-10-089-273-7

Perfect score: 107

Sequence: 1 TPEKEPTAAPPEAPEAPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 107   | 100.0       | 150    | 1     | ALL5 HEVBR         |
| 2          | 66    | 61.7        | 205    | 2     | Q39967 hevea bras  |
| 3          | 64.5  | 60.3        | 414    | 2     | Q75A16 ashbya goss |
| 4          | 64    | 59.8        | 3410   | 2     | Q6E102 streptococc |
| 5          | 63    | 58.9        | 454    | 2     | Q7N000 rattus norv |
| 6          | 63    | 58.9        | 834    | 2     | Q8WZ50 homo sapien |
| 7          | 63    | 58.9        | 7962   | 2     | Q10465 homo sapien |
| 8          | 63    | 58.9        | 34350  | 2     | Q8WZ42 homo sapien |
| 9          | 61.5  | 57.5        | 501    | 2     | Q9L714 caulobacter |
| 10         | 61.5  | 57.5        | 501    | 2     | Q9A496 caulobacter |
| 11         | 61    | 57.0        | 286    | 2     | Q6AE16 leifsonia x |
| 12         | 61    | 57.0        | 329    | 2     | Q7X1G9 oryza sativ |
| 13         | 61    | 57.0        | 666    | 2     | Q69740 mycobacteri |
| 14         | 61    | 57.0        | 666    | 2     | Q7TVG2 mycobacteri |
| 15         | 61    | 57.0        | 1581   | 2     | Q6PDJ2 mus musculu |
| 16         | 61    | 57.0        | 1848   | 2     | Q8FZM8 mus musculu |
| 17         | 60.5  | 56.5        | 444    | 2     | Q73VM4 mycobacteri |
| 18         | 60.5  | 56.5        | 1135   | 2     | Q812E1 mus musculu |
| 19         | 59    | 55.1        | 166    | 2     | Q9UVD0 pneumocysti |
| 20         | 59    | 55.1        | 221    | 1     | GRP1 STRAW         |
| 21         | 59    | 55.1        | 261    | 1     | OLG1 RAT           |
| 22         | 59    | 55.1        | 355    | 2     | Q8PC04 xanthomonas |
| 23         | 59    | 55.1        | 382    | 2     | Q84HN9 streptomyce |
| 24         | 59    | 55.1        | 856    | 2     | Q8QUT7 infectious  |
| 25         | 58    | 54.2        | 386    | 2     | Q9FPG5 chlamydomon |
| 26         | 58    | 54.2        | 411    | 2     | Q26766 trypanosoma |
| 27         | 58    | 54.2        | 439    | 1     | XP2_XENLA          |
| 28         | 57.5  | 53.7        | 462    | 2     | Q91U09 xenopus lae |
| 29         | 57    | 53.3        | 52     | 2     | Q6SPW9 plasmod pip |
| 30         | 57    | 53.3        | 59     | 2     | Q6SPW8 human immun |
| 31         | 57    | 53.3        | 78     | 2     | Q70VS1 human immun |

|    |    |      |     |   |                     |
|----|----|------|-----|---|---------------------|
| 32 | 57 | 53.3 | 117 | 2 | Q857A8 mycobacteri  |
| 33 | 57 | 53.3 | 176 | 2 | Q6R1F9 human immun  |
| 34 | 57 | 53.3 | 225 | 2 | Q6CAA2 yarrowia li  |
| 35 | 57 | 53.3 | 304 | 2 | Q9N3D1 caenorhabdi  |
| 36 | 57 | 53.3 | 326 | 2 | Q89JN5 bradyrhizob  |
| 37 | 57 | 53.3 | 361 | 2 | Q8YH56 bruceella me |
| 38 | 57 | 53.3 | 379 | 2 | Q99XV2 streptococc  |
| 39 | 57 | 53.3 | 410 | 2 | Q8G0Q0 bruceella su |
| 40 | 57 | 53.3 | 447 | 2 | Q6C5L8 yarrowia li  |
| 41 | 57 | 53.3 | 488 | 2 | Q80363 human immun  |
| 42 | 57 | 53.3 | 497 | 2 | Q90MF8 human immun  |
| 43 | 57 | 53.3 | 497 | 2 | Q90MF9 human immun  |
| 44 | 57 | 53.3 | 497 | 2 | Q8AC88 human immun  |
| 45 | 57 | 53.3 | 498 | 2 | Q6X6Y5 human immun  |

## ALIGNMENTS

RESULT 1

ALL5 HEVBR STANDARD; PRT; 150 AA.

AC Q39967; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Major latex allergen Hev b 5.

OS Hevea brasiliensis (Para rubber tree).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Microandreae;

OC Hevea.

OX NCBI\_TaxID=3981;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. RRIM 600; TISSUE=Latex;

RX MEDLINE=96411753; PubMed=8810306; DOI=10.1074/jbc.271.41.25389;

RA Akasawa A., Heieh L.-S., Martin B.M., Liu T., Lin Y.;

RT "Identification, cloning, and sequence of a major allergen (Hev b 5) from natural rubber latex (Hevea brasiliensis).";

RL J. Biol. Chem. 271:25394-25399(1996).

RN [2]

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND MASS SPECTROMETRY.

RC STRAIN=cv. RRIM 600; TISSUE=Latex;

RX MEDLINE=96411752; PubMed=8810305; DOI=10.1074/jbc.271.41.25389;

RA Akasawa A., Heieh L.-S., Martin B.M., Liu T., Lin Y.;

RT "A novel acidic allergen, Hev b 5, in latex. Purification, cloning and characterization.";

RL J. Biol. Chem. 271:25389-25393(1996).

CC -I- PTM: The N-terminus is blocked.

CC -I- MASS SPECTROMETRY: MW=16001.2; METHOD=MALDI; RANGE=1-150;

CC NOTE=Ref.2

CC -I- ALLERGEN: Causes an allergic reaction in human. Major latex allergen, a major cause of anaphylaxis in susceptible individuals, especially health care workers. 92% of health care workers with latex allergy have IgE specific to the Hev b 5 protein.

CC -I- SIMILARITY: To kiwi fruit protein PKIWI501.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC EMBL; U42640; AAC49447.1; --

CC EMBL; U51631; AAC49448.1; --

CC PIR; T10768; T10768.

CC Acetylation; Allergen; Direct protein sequencing.

FT INIT\_MET 0 0 Potential.

FT MOD\_RES 1 1 N-acetylalanine (Potential).

FT SEQUENCE 150 AA; 15959 MW; B37A8673A0A5F6B0 CRC64;

```
Query Match      100.0%; Score 107; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 9.6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPAPE 20
    |||||
Db 45 TPEKEPTAAPAEPAPE 64

RESULT 2
Q75AJ6 PRELIMINARY; PRT; 205 AA.
AC Q75AJ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ADL069WP.
GN ORFNames=ADL069W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegelé S.E., Brachat S., Dietrich F.S., Lerch A., Gaffney T.,
  Philippsen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016817; AAS51851.1; -.
DR AGD; ADL069W; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR003034; SAP.
DR PFAM; PF02037; SAP; 1.
DR SMART; SM00513; SAP; 1.
DR PROSITE; PS50800; SAP; 1.
SQ SEQUENCE 205 AA; 22228 MW; 15C6C867E3F198C3 CRC64;

Query Match      61.7%; Score 66; DB 2; Length 205;
Best Local Similarity 65.0%; Pred. No. 5.6;
Matches 13; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPAPE 20
    |||||
Db 89 TPEAEPAATPAEAPVAE 108

RESULT 3
Q6EI02 PRELIMINARY; PRT; 414 AA.
AC Q6EI02;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Collagen-like protein 3 (Fragment).
GN Names=Sc13B;
OS Streptococcus equi.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1336;
RN [1]
RP SEQUENCE FROM N.A.
RA Bujnicki J.M., Lukowski S., Kurowski M.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
  an amide bond (By similarity).
DR EMBL; AV326315; AAQ91576.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001899; Gram_pos_anchor.
DR PFAM; PF01391; Collagen; 3.
```

```
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PRO0015; GPOSANCHOR.
DR TIGRAME; TIGR01167; LPYTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Collagen; Peptidoglycan-anchor.
FT NON_TER 414
SQ SEQUENCE 414 AA; 42306 MW; ECB6A9A5140EA6FF CRC64;

Query Match      60.3%; Score 64.5; DB 2; Length 414;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 14; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 TPEKEPTAAPAEPAPE 20
    |||||
Db 323 TPEKKEQPAKPEPTPAPE 343

RESULT 4
Q7TN00 PRELIMINARY; PRT; 3410 AA.
AC Q7TN00;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiac titin N2BA isoform (Fragment).
GN Name=Ttn;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Ventricle;
RX MEDLINE=22669966; PubMed=12785098; DOI=10.1023/A:1023410523184;
RA Greaser M.L., Berri M., Warren C.M., Mozdziaik P.E.;
RT "Species variations in cDNA sequence and exon splicing patterns in the
  extensible I-band region of cardiac titin: relation to passive
  tension.";
RT J. Muscle Res. Cell Motil. 23:473-482(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Ventricle;
RA Greaser M.L., Berri M., Warren C.M., Mozdziaik P.E.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF525411; AAP80789.2; -.
DR HSSP; P13596; IQZ1.
DR GO; GO:0008307; F:structural constituent of muscle; IEA.
DR GO; GO:007517; P:muscle development; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003596; Ig_V.
DR InterPro; IPR004168; PPAK_motif.
DR Pfam; PF02818; PPAK; 7.
DR SMART; SM00409; IG; 29.
DR SMART; SM00408; IGC2; 28.
DR SMART; SM00406; IGV; 8.
DR PROSITE; PS50835; IG_LIKE; 28.
FT NON_TER 1
FT NON_TER 3410
SQ SEQUENCE 3410 AA; 375757 MW; 81BC913938B1E491 CRC64;

Query Match      59.8%; Score 64; DB 2; Length 3410;
Best Local Similarity 55.0%; Pred. No. 1.6e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPAPE 20
    |||||
Db 3178 TPEKVPVVPKKPEAPPK 3197

RESULT 5
Q8WZ50
```



DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Titin.  
 GN Name=TTN;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20309627; PubMed=10850961;  
 RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereuse F.,  
 RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,  
 RA Granzier H., Labeit S.;  
 RT "Series of exon-skipping events in the elastic spring region of titin  
 RT as the structural basis for myofibrillar elastic diversity.";  
 RL Circ. Res. 86:1114-1121(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21573839; PubMed=11717165;  
 RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,  
 RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,  
 RA Labeit S.;  
 RT "The complete gene sequence of titin, expression of an unusual -700  
 RT kDa titin isoform and its interaction with obscurin identify a novel  
 RT Z-line to I-band linking system.";  
 RL Circ. Res. 89:1065-1072(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Centner S.B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ277892; CAD12456.1; -.  
 DR PDB; 1TNN; NMR; @=33480-33579.  
 DR PDB; 1TNN; NMR; @=33480-33579.  
 DR Genew; HGNC:12403; TTN.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0004896; F:hematopoietin/interferon-class (D200-domain...); IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0008307; F:structural constituent of muscle; IEA.  
 DR GO; GO:0007517; P:muscle development; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR Pfam; PF00041; fn3; 132.  
 DR Pfam; PF00047; ig; 3.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF02818; PPAK; 53.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00060; FN3; 132.  
 DR SMART; SM00408; IGC2; 65.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00933; FGGV\_KINASES\_1; UNKNOWN\_1.  
 DR PROSITE; PS00853; FN3; 132.  
 DR PROSITE; PS00835; IG LIKE; UNKNOWN\_1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
 SQ SEQUENCE 34350 AA; 3816189 MW; 5B1120058A7CE58A CRC64;  
 Query Match 58.9%; Score 63; DB 2; Length 34350;  
 Best Local Similarity 57.9%; Pred. No. 2.1e+03;  
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 PEKKEPTAAPAEPAPE 20  
 Db 11489 PEKKVPEAIKKPKESPPE 11507  
 RESULT 9  
 ID Q9L714  
 AC Q9L714 PRELIMINARY; PRT; 501 AA.  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CpaF.  
 GN Name=cpaF;  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CB15;  
 RX MEDLINE=20341079; PubMed=10880436; DOI=10.1093/emboj/19.13.3223;  
 RA Skerker J.M., Shapiro L.;  
 RT "Identification and cell cycle control of a novel pilus system in  
 RT Caulobacter crescentus.";  
 RL EMBO J. 19:3223-3234(2000).  
 DR EMBL; AF229646; AAF40195.1; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR001482; GSP11\_E.  
 DR InterPro; IPR000217; Tubulin.  
 DR Pfam; PF00437; GSP11\_E; 1.  
 DR ProDom; PD000739; GSP11\_E; 1.  
 DR PROSITE; PS00227; TUBULIN; UNKNOWN\_1.  
 SQ SEQUENCE 501 AA; 54306 MW; 1F9DFBF14A04A609 CRC64;  
 Query Match 57.5%; Score 61.5; DB 2; Length 501;  
 Best Local Similarity 63.2%; Pred. No. 44;  
 Matches 12; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
 QY 2 PEKKEPTAAPAEPAPE 20  
 Db 30 PORVEFVAAP-EKAPAPK 47  
 RESULT 10  
 Q9A496  
 ID Q9A496 PRELIMINARY; PRT; 501 AA.  
 AC Q9A496;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Pilus assembly protein CpaF.  
 GN OrderedLocustNames=CC2942;  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;  
 RA Niemman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,  
 RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,  
 RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AF005958; AAK24904.1; -.  
 DR PIR; D87613; D87613.  
 DR TIGR; CC2942; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR001482; GSP11\_E.  
 DR InterPro; IPR000217; Tubulin.  
 DR Pfam; PF00437; GSP11\_E; 1.  
 DR ProDom; PD000739; GSP11\_E; 1.



```

DR PROSITE; PS00227; TUBULIN; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 501 AA; 54276 MW; 64EDB65A3C5DD7DA CRC64;

Query Match 57.5%; Score 61.5; DB 2; Length 501;
Best Local Similarity 63.2%; Pred. No. 44;
Matches 12; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPAEPEAPAP 20
DB 30 PQRVFVAAP-EKAPAPK 47

RESULT 11
Q6AE16 PRELIMINARY; PRT; 286 AA.
AC Q6AE16
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=Lxx15820;
OS Leifsonia xyl (subsp. xyl).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F. Jr., Carver H., Coutinho L.L., El-Dorri H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gagliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Teal S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT xyl subsp. xyl."
RL Mol. Plant Microbe Interact. 17:827-836 (2004).
DR EMBL; AE016822; AAT89380.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 286 AA; 29870 MW; 06384AD152D346DE CRC64;

Query Match 57.0%; Score 61; DB 2; Length 286;
Best Local Similarity 63.2%; Pred. No. 29;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPAP 19
DB 5 TDEPEEPPAASRHPEAPAP 23

RESULT 12
Q7XKLG9 PRELIMINARY; PRT; 329 AA.
AC Q7XKLG9
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSUNBA0039C07.7 protein.
GN Name=OSUNBA0039C07.7;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447439; DOI=10.1038/nature01183;

Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL731591; CAE05151.2; -.
DR HSSP; Q9RAG3; IQVR.
DR Gramene; Q7XKLG9; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR001005; Myb_DNA_binding.
DR PROSITE; PS50090; MYB_3; 1.
SQ SEQUENCE 329 AA; 34882 MW; 0078BF0DFCAC094E CRC64;

Query Match 57.0%; Score 61; DB 2; Length 329;
Best Local Similarity 63.2%; Pred. No. 33;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPAP 19
DB 195 TTPATTAAAPPPAPAP 213

RESULT 13
O69740 PRELIMINARY; PRT; 666 AA.
ID O69740
AC O69740; Q7DAP3;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CONSERVED HYPOTHETICAL PROLINE AND ALANINE RICH PROTEIN (Hypothetical
DE protein).
GN OrderedLocusNames=MT3990; Rv3876;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230; DOI=10.1038/311159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CDC 1551 / Oshkosh;
RC MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1126/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."

```

```

RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; BX842584; CAA17968.1; -.
DR EMBL; AB000516; AAK48358.1; -.
DR PIR; B70803; B70803.
DR TIGR; MT3990; -.
DR Tuberculist; RV3876; -.
KW Hypothetical protein.
SQ SEQUENCE 666 AA; 70644 MW; 752E072FB8B10111 CRC64;

Query Match 57.0%; Score 61; DB 2; Length 666;
Best Local Similarity 55.6%; Pred. No. 67;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEAPAP 19
DB 156 PTFQTGAPQQPESAP 173

RESULT 14
Q7TVG2 PRELIMINARY; PRT; 666 AA.
AC Q7TVG2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CONSERVED HYPOTHETICAL PROLINE AND ALANINE RICH PROTEIN.
GN OrderedLocustNames=Mb3906;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248347; CAD96092.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 666 AA; 70644 MW; 752E072FB8B10111 CRC64;

Query Match 57.0%; Score 61; DB 2; Length 666;
Best Local Similarity 55.6%; Pred. No. 67;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEAPAP 19
DB 156 PTFQTGAPQQPESAP 173

RESULT 15
Q6PDJ2 PRELIMINARY; PRT; 1581 AA.
AC Q6PDJ2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BC058674 protein (fragment).
GN Name=BC058674;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

```

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC058674; AAH58674.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000637; A+T hook.
DR Pfam; PF02178; AT_hook; 2.
FT NON_TER 1
SQ SEQUENCE 1581 AA; 166617 MW; 82F3018F6EAEFBA6 CRC64;

Query Match 57.0%; Score 61; DB 2; Length 1581;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 14; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

QY 1 TPEK-----EPTAAPAEPEAPE 20
DB 1245 TPEKTSKPPPEAPEPAVEPPAPE 1270

RESULT 16
Q69ZN8 PRELIMINARY; PRT; 1848 AA.
AC Q69ZN8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIAA1205 protein (Fragment).
GN Name=MKIAA1205;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic tail;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173130; BAD32408.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000637; A+T hook.
DR Pfam; PF02178; AT_hook; 2.
FT NON_TER 1
SQ SEQUENCE 1848 AA; 192477 MW; 51696D30321FF565 CRC64;

```

```
Query Match      57.0%; Score 61; DB 2; Length 1848;
Best Local Similarity 53.8%; Pred. No. 1.9e+02;
Matches 14; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

QY 1 TPEK-----EPTAAPAPAPAP 20
    ||||| ||||| ||||| ||||| |||||
Db 1512 TPEKMTSEKPPAPAPAPAPAP 1537

RESULT 17
Q73VM4 PRELIMINARY; PRT; 444 AA.
AC Q73VM4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FtsY.
GN Names=ftsY; OrderedLocusNames=MAP2989c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017237; AA505306.1; -.
DR GO; GO:0005786; C:signal recognition particle (sensu Eukaryota); IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0001666; F:nucleotide binding; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0006514; P:SRP-dependent cotranslational protein-membr. . .; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004390; ABC_transp_FtsY.
DR InterPro; IPR000897; SRP54.
DR Pfam; PF00448; SRP54; 1.
DR Pfam; PF02881; SRP54_N; 1.
DR ProDom; PD000819; SRP54; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00664; ftsY; 1.
DR PROSITE; PS00300; SRP54; 1.
KW Complete proteome.
SQ SEQUENCE 444 AA; 46236 MW; 20D1AADB9F02890 CRC64;

Query Match      56.5%; Score 60.5; DB 2; Length 444;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 14; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

QY 2 PEKEPTA-----APAPAPAP 20
    ||||| ||||| ||||| ||||| |||||
Db 102 PEPETPAAPAPAPAPAPAP 129

RESULT 18
Q812E1 PRELIMINARY; PRT; 1135 AA.
AC Q812E1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Connectin/titin N2A-PEVK (Fragment).
GN Name=Ttn;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mdm;
RA Sorimachi H., Labelit S.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB100271; BAC55524.2; -.
DR HSSP; PS6276; 1FHG.
DR MGD; MGI:98864; Ttn.
DR GO; GO:0005859; C:muscle myosin; TAS.
DR GO; GO:0030017; C:sarcomere; IDA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
DR GO; GO:0006938; P:sarcomere alignment; TAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR004168; PPAK_motif.
DR Pfam; PF02818; PPAK; 21.
DR SMART; SM0408; IGC2; 1.
DR PROSITE; PS0835; IG_LIKE; 2.
FT NON TER 1
FT NON TER 1135
SQ SEQUENCE 1135 AA; 130442 MW; 30647A8BDCA51BAE CRC64;

Query Match      56.5%; Score 60.5; DB 2; Length 1135;
Best Local Similarity 59.1%; Pred. No. 1.3e+02;
Matches 13; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 2 PEKEPTAAPAPAPAPAP 20
    ||||| ||||| ||||| ||||| |||||
Db 591 PPKKEPAPAPAPAPAP 612

RESULT 19
Q9UVD0 PRELIMINARY; PRT; 166 AA.
AC Q9UVD0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kexin-like serine endoprotease (Fragment).
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rat;
RA Russian D.A., Andrawis-Sorial V., Kovacs J.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009223; AAF21602.1; -.
DR MEROPS; S08.011; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR03882; FtsII_extensin.
DR PRINTS; PR01218; PSTLEXTENSIN.
KW Protease.
FT NON TER 1
FT NON TER 166
SQ SEQUENCE 166 AA; 17673 MW; C6547D30CCA3CA05 CRC64;

Query Match      55.1%; Score 59; DB 2; Length 166;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAPAPAP 19
    ||||| ||||| ||||| |||||
Db 42 PPQPEPPAPAPAP 59

RESULT 20
GRP1_STRAW STANDARD; PRT; 221 AA.
AC Q82EX8;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE GrpE protein 1 (HSP-70 cofactor 1).
GN Name=grpE1; OrderedLocusNames=SAV4485;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
```



```

RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.B., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Camavari F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sana J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012193; AAM40261.1; -.
DR InterPro; IPR007730; SPOR.
DR Pfam; PF05036; SPOR; 4.
KW Complete proteome.
SQ SEQUENCE 355 AA; 35513 MW; 763EAB648F4339D CRC64;

Query Match 55.1%; Score 59; DB 2; Length 355;
Best Local Similarity 68.8%; Pred. No. 60;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 EEPTAAPAEPEAPE 20
Db 239 EAPRPAPAKPEAPKE 254

RESULT 23
ID Q84HN9 PRELIMINARY; PRT; 382 AA.
AC Q84HN9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Unbu.
GN Name=unbu;
OS Streptomyces ghanaensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=35758;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-12104;
RX MEDLINE=22447897; PubMed=12536216;
RA Zazopoulos E., Huang K., Scaffa A., Liu W., Bachmann B.O., Nonaka K.,
RA Ahlert J., Thorson J.S., Shen B., Farnet C.M.;
RT "A genomics-guided approach for discovering and expressing cryptic
RT metabolic pathways.";
RL Nat. Biotechnol. 21:187-190(2003).
DR EMBL; AP546141; AAC25843.1; -.
DR NCBI_TaxID=35758;
SQ SEQUENCE 382 AA; 40391 MW; 5A13A3F8006410DA CRC64;

Query Match 55.1%; Score 59; DB 2; Length 382;
Best Local Similarity 61.1%; Pred. No. 64;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PEKEPTAAPAEPEAP 19
Db 318 PSGAQPDAPAEAPAP 335

RESULT 24
ID Q8QUT7 PRELIMINARY; PRT; 856 AA.
AC Q8QUT7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ORF023R.
OS Infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae;
OC unclassified Iridoviridae.
OX NCBI_TaxID=180170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus.";
RL Virology 291:126-139(2001).
DR EMBL; AF371960; AAL98747.1; -.
DR HSSP; P02468; INPE.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR009030; Grow_fac_recept.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00053; Laminin_EGF; 6.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00180; EGF_Lam; 5.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 5.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 5.
SQ SEQUENCE 856 AA; 91952 MW; C15B95B7649530EB CRC64;

Query Match 55.1%; Score 59; DB 2; Length 856;
Best Local Similarity 62.5%; Pred. No. 14e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PEKEPTAAPAEPEAP 17
Db 712 PERAPPPAPSQPEAP 727

RESULT 25
ID Q9FPQ5 PRELIMINARY; PRT; 386 AA.
AC Q9FPQ5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gamete-specific hydroxyproline-rich glycoprotein a2.
OS Chlamydomonas reinhardtii.
OC Rukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-621;
RX MEDLINE=21663714; PubMed=11805055;
RA Ferris P.J., Armbrust E.W., Goodenough U.W.;
RT "Genetic structure of the mating-type locus of Chlamydomonas
RT reinhardtii.";
RL Genetics 160:181-200(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-621;
RX MEDLINE=21159092; PubMed=11258910; DOI=10.1021/bi0023605;
RA Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
RA Goodenough U.W.;
RT "Glycosylated polyproline II rods with kinks as a structural motif in
RT plant hydroxyproline-rich glycoproteins.";
RL Biochemistry 40:2978-2987(2001).
DR EMBL; AF309495; AAG45421.1; -.
DR NCBI_TaxID=3055;
SQ SEQUENCE 386 AA; 38775 MW; 7B4F4CA7D95C2EF7 CRC64;

Query Match 54.2%; Score 58; DB 2; Length 386;
Best Local Similarity 52.6%; Pred. No. 84;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```
OY 2 PEKEEPTAAPAEPEAPE 20
Db 223 PSEPPSPAPPSPPEPPSP 241

RESULT 26
O26766
ID Q26766 PRELIMINARY; PRT; 411 AA.
AC Q26766;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Flagellar antigen (Fragment).
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=stock TREU 1285;
RX MEDLINE=95241149; PubMed=7724233;
RA Imboden M., Mueller N., Hemphill A., Mattioli, Seebeck T.;
RT "Repetitive proteins from the flagellar cytoskeleton of African
RL Trypanosomes are diagnostically useful antigens.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=stock TREU 1285;
RA Michael I.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z36281; CAA85287.1; --
DR PIR; S47436; S47436.
DR GO; GO:0019861; C:flagellum; IEA.
DR InterPro; IPR009050; Globin_like.
KW Flagellum.
FT NON TER 1
SQ SEQUENCE 411 AA; 43321 MW; 4614A0EA370D180E CRC64;

Query Match 54.2%; Score 58; DB 2; Length 411;
Best Local Similarity 70.0%; Pred. NO. 89;
Matches 14; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

OY 3 EKEEPTAAPAE--PEAPAE 20
Db 378 ELEEPQAPAEAPQPEAAPE 397

RESULT 27
XP2_XENLA
ID XP2_XENLA STANDARD; PRT; 439 AA.
AC F17437; Q08944;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Skin secretory protein xp2 precursor (APEG protein).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE OF 1-25 AND 344-439 FROM N.A. (ISOFORM 2).
RX MEDLINE=92332564; PubMed=1629230;
RA Hauser F., Roeben C., Hoffmann W.;
RT "xp2, a new member of the P-domain peptide family of potential growth
factors, is synthesized in Xenopus laevis skin.";
RL J. Biol. Chem. 267:14451-14455(1992).
RN [2]
RP SEQUENCE OF 3-439 FROM N.A. (ISOFORM 1).
RC TISSUE=Skin;
RX MEDLINE=90127399; PubMed=2298293; DOI=10.1016/0014-5793(90)80088-Z;
RA Gmachl M., Berger H., Thalhammer J., Kreil G.;
RT "Dermal glands of Xenopus laevis contain a highly

RT repetitive amino acid sequence. ";
RL FEBS Lett. 260:145-148(1990).
CC -!- FUNCTION: May act as a growth factor in the germinal layer of the
epidermis. May also be involved in growth of regenerating glands
and in protection of the skin from the external environment.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=APEG;
CC IsoId=P17437-1; Sequence=Displayed;
CC Name=2; Synonyms=XP2;
CC IsoId=P17437-2; Sequence=VSP_004652;
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Contains 2 P-type (trefoil) domains.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
frameshift in position 392.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90095; AAA50001.1; -
DR EMBL; X51394; CAA35759.1; ALT_FRAME.
DR PIR; A37331; A37331.
DR PIR; S07498; SKXLAG.
DR HSSP; P04155; 1P32.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil_2.
DR PRINTS; PRO0680; PTREFOIL.
DR PROSITE; PS00025; P_TREFOIL; 2.
KW Alternative splicing; Growth factor; Pyrrolidone carboxylic acid;
Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 439 Skin secretory protein xp2.
FT DOMAIN 26 343 33 X approximate repeats of G-G(0,1)-
[EV] (0,1)-A-P-[A-P] (1,3)-A-E.
FT REPEAT 26 33 1.
FT REPEAT 34 41 2.
FT REPEAT 42 51 3.
FT REPEAT 52 59 4.
FT REPEAT 60 69 5.
FT REPEAT 70 77 6 (approximate).
FT REPEAT 78 87 7.
FT REPEAT 88 97 8.
FT REPEAT 98 107 9.
FT REPEAT 108 115 10.
FT REPEAT 116 125 11.
FT REPEAT 126 135 12.
FT REPEAT 136 145 13.
FT REPEAT 146 153 14.
FT REPEAT 154 163 15.
FT REPEAT 164 173 16 (approximate).
FT REPEAT 174 183 17.
FT REPEAT 184 193 18.
FT REPEAT 194 203 19.
FT REPEAT 204 215 20.
FT REPEAT 216 225 21.
FT REPEAT 226 235 22.
FT REPEAT 236 245 23.
FT REPEAT 246 255 24.
FT REPEAT 256 265 25.
FT REPEAT 266 275 26.
FT REPEAT 276 285 27.
FT REPEAT 286 293 28.
FT REPEAT 294 303 29.
FT REPEAT 304 313 30.
FT REPEAT 314 321 31 (approximate).
FT REPEAT 322 331 32 (approximate).
FT REPEAT 332 343 33 (approximate).
```

```

FT DOMAIN 350 391 P-type 1.
FT DOMAIN 397 438 P-type 2.
FT MOD RES 23 23 Pyrolidone carboxylic acid (Probable).
FT DISULFID 351 377 By similarity.
FT DISULFID 361 376 By similarity.
FT DISULFID 371 388 By similarity.
FT DISULFID 398 424 By similarity.
FT DISULFID 408 423 By similarity.
FT DISULFID 418 435 By similarity.
FT VARSPPLIC 26 343 Missing (in isoform 2).
FT CONFLICT 3 3 /FTID=VSP_004652.
FT CONFLICT 18 18 H -> S (in Ref. 2).
FT CONFLICT 18 18 C -> W (in Ref. 2).
SQ SEQUENCE 439 AA; 41173 MW; 38C4A4B57CBAE778 CRC64;

Query Match 54.2%; Score 58; DB 1; Length 439;
Best Local Similarity 70.6%; Pred. No. 96;
Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKEPTAAPAEPEAPAP 19
Db 125 EGEAPAPAEGEAPAP 141

RESULT 28
Q91U09 PRELIMINARY; PRT; 462 AA.
AC Q91U09
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RepA protein.
GN Name=repA;
OS Plasmid pIPO2T.
OC other sequences; broad host range plasmids.
OX NCBI_TaxID=141833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22051032; PubMed=12055285;
RA Tsuchi A., Schneider S., Selbitschka W., Puhler A., van Overbeek L.S.,
RA Smalla K., Thomas C.M., Bailey M.J., Forney L.J., Weightman A.,
RA Ceglowski P., Pembroke T., Tietze E., Schroeder G., Lanka E.,
RA van Elsas J.D.;
RT "The complete nucleotide sequence and environmental distribution of
RT the cryptic, conjugative, broad-host-range plasmid pIPO2 isolated from
RT bacteria of the wheat rhizosphere.";
RL Microbiology 148:1637-1653(2002).
DR EMBL; AJ297913; CAC82776.1; -.
DR Pfam; PF04796; RepA_C; 1.
SQ SEQUENCE 462 AA; 50710 MW; F29FF78299DC0CF0 CRC64;

Query Match 53.7%; Score 57.5; DB 2; Length 462;
Best Local Similarity 65.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 2 PEKEPT---AAPAEPEAPA 18
Db 115 PEAAKPTQAPAAPEPEKPA 134

RESULT 29
Q6SPW9 PRELIMINARY; PRT; 52 AA.
AC Q6SPW9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gag protein (Fragment).
GN Name=gag;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Lastere S., Dalban C., Collin G., Descamps D., Girard P.-M.,
RA Clavel F., Costagliola D., Brun-Vezinet F.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453181; AAR15064.1; -.
FT NON_TER 1 1
SQ SEQUENCE 52 AA; 5500 MW; BA7289CC52AC8260 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 52;
Best Local Similarity 64.7%; Pred. No. 15;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 KEPTAAPAEPEAPAE 20
Db 4 RPEPTAPAEPTAPAE 20

RESULT 30
Q6SPW8 PRELIMINARY; PRT; 59 AA.
AC Q6SPW8
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gag protein (Fragment).
GN Name=gag;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lastere S., Dalban C., Collin G., Descamps D., Girard P.-M.,
RA Clavel F., Costagliola D., Brun-Vezinet F.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453182; AAR15065.1; -.
FT NON_TER 1 1
SQ SEQUENCE 59 AA; 6553 MW; 4D8522658E09927C CRC64;

Query Match 53.3%; Score 57; DB 2; Length 59;
Best Local Similarity 64.7%; Pred. No. 17;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 KEPTAAPAEPEAPAE 20
Db 4 RPEPTAPAEPTAPAE 20

RESULT 31
Q70VS1 PRELIMINARY; PRT; 78 AA.
AC Q70VS1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gag polyprotein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Esteves A., Santos M., Parreira R., Piedade J., Venenno T.,
RA Canas-Perreira W.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ534999; CAD59296.1; -.
KW Polypeptid.
FT NON_TER 1 1 protease.
FT CHAIN 69 78 P6 protein.
FT CHAIN <1 78
SQ SEQUENCE 78 AA; 8433 MW; 068180EF3883D10 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 78;
Best Local Similarity 64.7%; Pred. No. 22;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```
QY 4 KEEPTAAPAEPEAPE 20
SQ SEQUENCE 176 AA; 19207 MW; 3C303E3F0BDA3592 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 176;
Best Local Similarity 64.7%; Pred. No. 50;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 32
Q857A8 PRELIMINARY; PRT; 117 AA.
AC Q857A8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Gp26.
OS Mycobacteriophage Rosebush.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205874;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AX129334; AN01868.1; -.
SQ SEQUENCE 117 AA; 13055 MW; 3795BD4E2F1A1562 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 117;
Best Local Similarity 54.2%; Pred. No. 33;
Matches 13; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 TPEKEE---PTAAPAEPEAPE 20
DB 50 TPEPERLEAPQVPSTPEAVAPK 73

RESULT 33
Q6R1F9 PRELIMINARY; PRT; 176 AA.
AC Q6R1F9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Gag-pol polyprotein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RG AIEDRP;
RA Leigh Brown A.J., Frost S.D.W., Good B., Daar E.S., Simon V.,
RA Markowitz M., Collier A.C., Connick E., Conway B., Margolick J.B.,
RA Routhy J.-P., Corbell J., Hellmann N.S., Richman D.D., Little S.J.;
RT "Genetic Basis of Hypersusceptibility to Protease Inhibitors and Low
RT Replicative Capacity of Human Immunodeficiency Virus Type 1 Strains in
RT Primary Infection.";
RL J. Virol. 0:0-0(2004).
CC -1- SIMILARITY: Belongs to peptidase family A2.
DR EMBL; AY518996; AA521943.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_Aspartic.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Polypeptide; Polyprotein; Protease.
FT NON_TER 1 1
```



|     |  |
|-----|--|
| OC  | Rhabditidae; Peloderinae; Caenorhabditis.                      |
| OX  | NCBI_TaxID=6239;   |
| [1] |  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | MEDLINE=99069613; PubMed=9851916;                              |
| RX  | WormBase Consortium;   |
| RT  | "Genome sequence of the nematode C. elegans: a platform for    |
| RG  | investigating biology. The C. elegans Sequencing Consortium."; |
| RL  | Science 282:2012-2018(1998).                                   |
| RL  | [2]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Tin-Wollam A., Becker M., Graves T.;                           |
| RT  | "The sequence of C. elegans cosmid Y54E10BR.";                 |
| RG  | Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.        |
| RL  | [3]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.H.;  |
| RA  | Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.        |
| RL  | [4]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.        |
| RL  | [5]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.        |
| RL  | [6]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.        |
| RL  | [7]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.        |
| RL  | [8]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.        |
| RL  | [9]  |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Waterston R.;  |
| RA  | Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.        |
| RL  | [10]   |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Wilson R.;   |
| RA  | Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.        |
| RL  | [11]   |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | Wilson R.;   |
| RA  | Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.        |
| RL  | [12]   |
| RN  | SEQUENCE FROM N.A.   |
| RP  | STRAIN=Bristol N2;   |
| RC  | WormBase Consortium;   |
| RG  | Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.        |
| RL  | 1- SWIMLARIty: Contains 1 RING-type zinc finger.               |
| CC  | EMBL: AC024812; AAF59555.1; "                                  |
| DR  | WormBase; WBGene00021842; Y54E10BR.3.                          |
| DR  | WormPepp; Y54E10BR.3; CE22470.                                 |
| DR  | GO:0000151; C:ubiquitin ligase complex; IEA.                   |
| DR  | GO:00004842; F:ubiquitin-protein ligase activity; IEA.         |
| DR  | GO:0008270; F:zinc ion binding; IEA.                           |

```

GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; zfi-C3HC4; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 304 AA; 33447 MW; 3E522361AAP54C44 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 304;
Best Local Similarity 47.4%; Pred. No. 86;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEPEAPE 20
Db 78 PQQQPQGAPQEPETQAE 96

RESULT 36
Q89JN5 PRELIMINARY; PRT; 326 AA.
ID Q89JN5
AC Q89JN5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bli5248 protein.
DE OrderedLocusNames=bli5248;
GN Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OC NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisaawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL: AP005954; BAC50513.1; -.
KW Complete proteome.
SQ SEQUENCE 326 AA; 34942 MW; 8B795DCC41A1D444 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 326;
Best Local Similarity 45.0%; Pred. No. 92;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPE 20
Db 128 TPQREATAQPQQPQPQ 147

RESULT 37
Q8YH56 PRELIMINARY; PRT; 361 AA.
ID Q8YH56
AC Q8YH56;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein BMEI0948.
DE OrderedLocusNames=BMEI0948;
GN Brucella melitensis.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Muehler C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,

```

```

RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AB009535; AAL52129.1; -.
DR PIR; AF3370; AF3370.
KW Complete proteome.
SQ SEQUENCE 361 AA; 38894 MW; EB45E24903C61EE8 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 361;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 12; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 2 PEKEPTAAPAEPEAPAP 19
DB 229 PERE--SAAPAKPAAPAP 244
||:| :|||:|:|:|

RESULT 38
ID Q99XV2 PRELIMINARY; PRT; 379 AA.
AC Q99XV2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein SPY2009.
GN OrderedLocusNames=SPY2009;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=211926296; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AB006623; AAK34690.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Hypothetical protein;
KW Peptidoglycan-anchor.
SQ SEQUENCE 379 AA; 40352 MW; FFF78BDA501F15C4 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 379;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20
DB 221 SPSEEPSVAAPSETPSPE 240
||:| :|||:|:|:|

RESULT 39
Q8G0Q0 PRELIMINARY; PRT; 410 AA.
ID Q8G0Q0
AC Q8G0Q0;

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BRI1038;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014405; AAN29959.1; -.
DR TIGR; BR1038; -.
KW Complete proteome.
SQ SEQUENCE 410 AA; 44415 MW; 47E4A34F25E2FDF4 CRC64;

Query Match 53.3%; Score 57; DB 2; Length 410;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 12; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 2 PEKEPTAAPAEPEAPAP 19
DB 286 PERE--SAAPAKPAAPAP 301
||:| :|||:|:|:|

RESULT 40
Q8C5L8 PRELIMINARY; PRT; 447 AA.
ID Q6C5L8
AC Q6C5L8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome E of strain CLIB99 of Yarrowia
DE lipolytica (Fragment).
GN ORFNames=YAL10E16929g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E., Barbe
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Bagnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boirame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenné S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN=CLIB99;  
RA Genoscope;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
CC -|- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.  
CC -|- SIMILARITY: Contains 1 lipoyl-binding domain.  
DR EMBL; CR382131; CAG79637.1; -.  
DR GO; GO:0045252; C:oxoglutarate dehydrogenase complex; IEA.  
DR GO; GO:0008415; F:acyltransferase activity; IEA.  
DR GO; GO:0004149; F:dihydrolipoyllysine-residue succinyltransferase. . .; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.  
DR InterPro; IPR001078; 2Oxoacid\_dh.  
DR InterPro; IPR000089; Biotin\_lipoyl.  
DR InterPro; IPR011053; Hybrid\_motif.  
DR InterPro; IPR003016; Lipoyl\_BS.  
DR InterPro; IPR006255; SucB.  
DR Pfam; PF00198; 2-oxoacid\_dh; 1.  
DR Pfam; PF00364; Biotin\_lipoyl; 1.  
DR ProDom; PD001115; 2Oxoacid\_dh; 1.  
DR TIGRFAMs; TIGR01347; sucB; 1.  
DR PROSITE; PS00189; LIPOYL; 1.  
KW Acyltransferase; Lipoyl; Transferase.  
FT NON\_TER 1  
SQ SEQUENCE 447 AA; 48823 MW; 1DCBDC85516CA8B CRC64;  
  
Query Match 53.3%; Score 57; DB 2; Length 447;  
Best Local Similarity 65.0%; Pred.No. 1.3e+02;  
Matches 13; Conservative 1; Mismatches 4; Indels 2; Gaps 1;  
  
QY 1 TPEKEEPTAAPPEAPEAPE 20  
Db 157 TEEKEEP--APKEESAPAPK 174

Search completed: August 11, 2005, 09:57:46  
Job time : 175 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2005, 09:48:07 ; Search time 39 Seconds  
(without alignments)  
49.342 Million cell updates/sec

Title: US-10-089-273-7

Perfect score: 107

Sequence: 1 TPEKEPTAAPAEPEAPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length   | DB ID    | Description          |
|------------|-------|-------------|----------|----------|----------------------|
| 1          | 107   | 100.0       | 151      | 2 T10768 | latex allergen hev   |
| 2          | 63    | 58.9        | 7962     | 2 I38346 | elastic titin - hu   |
| 3          | 61.5  | 50.1        | 2 D87613 | 2        | pilus assembly prote |
| 4          | 61    | 57.0        | 666      | 2 B70803 | hypothetical prote   |
| 5          | 58    | 54.2        | 411      | 2 S47436 | flagellar antigen    |
| 6          | 58    | 54.2        | 416      | 1 SKXLG  | dermal gland prote   |
| 7          | 57    | 53.3        | 361      | 2 AF3370 | hypothetical prote   |
| 8          | 56.5  | 52.8        | 452      | 2 T32996 | hypothetical prote   |
| 9          | 56    | 52.3        | 2187     | 2 T30826 | neurofilament medi   |
| 10         | 55    | 51.4        | 801      | 2 T29018 | nascent polypeptid   |
| 11         | 55    | 51.4        | 1201     | 2 T29018 | hypothetical prote   |
| 12         | 55    | 51.4        | 3759     | 2 A35085 | hypothetical prote   |
| 13         | 55    | 51.4        | 3759     | 2 A35085 | trithorax protein    |
| 14         | 54    | 50.5        | 213      | 2 T44588 | probable transcrip   |
| 15         | 54    | 50.5        | 347      | 2 H75253 | hypothetical prote   |
| 16         | 54    | 50.5        | 365      | 2 B75398 | hypothetical prote   |
| 17         | 54    | 50.5        | 375      | 2 T08134 | oleosin-like prote   |
| 18         | 54    | 50.5        | 537      | 2 A46611 | myosin-binding pro   |
| 19         | 54    | 50.5        | 1870     | 2 S37671 | MHC class III hist   |
| 20         | 54    | 50.5        | 1872     | 2 S36152 | MHC class III hist   |
| 21         | 54    | 50.5        | 2142     | 2 B35098 | MHC class III hist   |
| 22         | 53.5  | 50.0        | 193      | 1 MORT4E | myosin alkali high   |
| 23         | 53.5  | 50.0        | 422      | 2 T24865 | hypothetical prote   |
| 24         | 53.5  | 50.0        | 455      | 2 E83598 | signal recognition   |
| 25         | 53    | 49.5        | 238      | 2 T26419 | hypothetical prote   |
| 26         | 53    | 49.5        | 409      | 2 T18726 | hypothetical prote   |
| 27         | 53    | 49.5        | 425      | 2 T18723 | hypothetical prote   |
| 28         | 53    | 49.5        | 544      | 2 T17547 | proline-rich prote   |
| 29         | 53    | 49.5        | 547      | 2 H83018 | dihydrolipoamide a   |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 53   | 49.5 | 839  | 2 F75518 | hypothetical prote |
| 31 | 53   | 49.5 | 1006 | 2 T42731 | atrophin-1 related |
| 32 | 53   | 49.5 | 1127 | 2 T32404 | hypothetical prote |
| 33 | 52   | 48.6 | 185  | 2 T00519 | proline-rich prote |
| 34 | 52   | 48.6 | 214  | 2 S29793 | protein ORF 214 (a |
| 35 | 52   | 48.6 | 237  | 2 T35351 | hypothetical prote |
| 36 | 52   | 48.6 | 536  | 2 H71563 | hypothetical prote |
| 37 | 52   | 48.6 | 574  | 2 T43556 | wiskott-Aldrich sy |
| 38 | 52   | 48.6 | 574  | 2 T38819 | wiskott-Aldrich sy |
| 39 | 52   | 48.6 | 646  | 2 G83146 | penicillin-binding |
| 40 | 52   | 48.6 | 647  | 2 T43952 | hypothetical prote |
| 41 | 52   | 48.6 | 655  | 2 AD2422 | DNA polymerase III |
| 42 | 52   | 48.6 | 759  | 2 T44142 | DRI protein [impor |
| 43 | 52   | 48.6 | 901  | 2 A49227 | sialidase - Actino |
| 44 | 52   | 48.6 | 1013 | 2 C83771 | hypothetical prote |
| 45 | 51.5 | 48.1 | 161  | 2 G97919 | acetyl-CoA carboxy |

ALIGNMENTS

RESULT 1

T10768

latex allergen hev b5 - Para rubber tree

C:Species: Hevea brasiliensis (Para rubber tree)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004

C:Accession: T10768

R:Slater, J.E.; Vedvick, T.; Arthur-Smith, A.; Trybul, D.E.; Kekwick, R.K.O.

J. Biol. Chem. 271, 25394-25399, 1996

A>Title: Identification, cloning and sequence of a major allergen (hev b 5) from natural

A:Reference number: Z17129; MUID:96411753; PMID:8810306

A:Accession: T10768

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-151 <SLA>

A:Cross-references: UNIPROT:Q39967; EMBL:U42640; NID:gl480456; PIDN:AAC49447.1; PID:gl480

A:Experimental source: cv. RRIM 600; latex

Query Match 100.0%; Score 107; DB 2; Length 151;

Best Local Similarity 100.0%; Pred. No. 1,1e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPEKEPTAAPAEPEAPE 20

Db 46 TPEKEPTAAPAEPEAPE 65

RESULT 2

I38346

elastic titin - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004

C:Accession: I38346

R:Labeit, S.; Kolmerer, B.

Science 270, 293-296, 1995

A>Title: Titins: Giant proteins in charge of muscle ultrastructure and elasticity.

A:Reference number: A57430; MUID:96026330; PMID:7569978

A:Accession: I38346

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-7962 <RES>

A:Cross-references: UNIPROT:Q10465; EMBL:X90569; NID:gl017426; PIDN:CAA62189.1; PID:gl01

C:Genetics:

A:Gene: GDB:TTN

A:Cross-references: GDB:127867; OMIM:188840

A:Map position: 2q31-2q31

Query Match 58.9%; Score 63; DB 2; Length 7962;

Best Local Similarity 57.9%; Pred. No. 54;

Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PEKEPTAAPAEPEAPE 20

|||: | | | : | | |



QY 2 PEKEEPTAAPAEPEAPAP 19  
Db 229 PERE--SAAPAKPAAPAP 244

RESULT 8  
T32996  
hypothetical protein W02F12.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T32996  
R;Du, Z.; Gattung, S.; Andrews, S.  
A;Title: Differential splicing converts alphaNAC into a muscle-specific DNA-binding active factor  
A;Description: The sequence of C. elegans cosmid W02F12.  
A;Reference number: Z21261  
A;Accession: T32996  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-452 <DUZ>  
A;Cross-references: UNIPROT:O45148; EMBL:AF047653; PIDN:AAC04462.1; GSPDB:GN00023; CESP:  
A;Experimental source: strain Bristol N2; clone W02F12  
C;Genetics:  
A;Gene: CESP:W02F12.5  
A;Map position: 5  
A;Introns: 19/3; 33/1; 60/1; 145/3; 269/2; 408/3  
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 52.8%; Score 56.5; DB 2; Length 452;  
Best Local Similarity 63.2%; Pred. No. 19;  
Matches 12; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPAEPEA-PAP 19  
Db 148 PAKEEPSAPAEKSKAP 166

RESULT 9  
I50479  
neurofilament medium protein - goldfish  
C;Species: Carassius auratus (goldfish)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 09-Jul-2004  
C;Accession: I50479  
R;Glasgow, E.; Hall, C.M.; Schechter, N.  
J. Neurochem. 63, 52-61, 1994  
A;Title: Organization, sequence, and expression of a gene encoding goldfish neurofilament  
A;Reference number: I50479; MUID:94267484; PMID:8207446  
A;Accession: I50479  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-798 <GLA>  
A;Cross-references: UNIPROT:Q90307; GB:L09742; NID:G213019; PIDN:AAA72341.1; PID:G213020  
C;Genetics:  
A;Introns: 343/3; 385/2  
C;Superfamily: cytoskeletal keratin

Query Match 52.3%; Score 56; DB 2; Length 798;  
Best Local Similarity 55.0%; Pred. No. 38;  
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPE 20  
Db 677 SPTKEEPAEPKPEAPE 696

RESULT 10  
T30826  
nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse  
N;Alternate names: alpha-NAC protein  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30826  
R;Yotov, W.V.; St-Arnaud, R.  
Genes Dev. 10, 1763-1772, 1996

A;Title: Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle-  
A;Reference number: Z20889; MUID:96312450; PMID:8698236  
A;Accession: T30826  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-2187 <YOT>  
A;Cross-references: UNIPROT:P70670; EMBL:U48363; NID:G1666688; PID:G1666689; PIDN:AAB187;  
C;Genetics:  
A;Gene: Naca  
A;Map position: 10  
A;Introns: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3  
A;Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding active  
C;Keywords: alternative splicing; DNA binding; transcription factor

Query Match 52.3%; Score 56; DB 2; Length 2187;  
Best Local Similarity 55.0%; Pred. No. 99;  
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPE 20  
Db 1351 SPRKAPKTAAPKETPAPSPE 1370

RESULT 11  
T29018  
hypothetical protein ZK84.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T29018  
R;Kirsten, J.  
submitted to the EMBL Data Library, April 1995  
A;Description: The sequence of C. elegans cosmid ZK84.  
A;Reference number: Z20553  
A;Accession: T29018  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-801 <KIR>  
A;Cross-references: UNIPROT:Q23635; EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:ZK84  
A;Experimental source: strain Bristol N2; clone ZK84  
C;Genetics:  
A;Gene: CESP:ZK84.1  
A;Map position: 2  
A;Introns: 22/2; 45/3; 108/1  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 51.4%; Score 55; DB 2; Length 801;  
Best Local Similarity 62.5%; Pred. No. 49;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 KEPEPTAAPAEPEAPAP 19  
Db 718 EEAPAPAPAPETPAP 733

RESULT 12  
T29329  
hypothetical protein F54D11.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T29329  
R;Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, July 1996  
A;Description: The sequence of C. elegans cosmid F54D11.  
A;Reference number: Z20606  
A;Accession: T29329  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1201 <PAU>  
A;Cross-references: UNIPROT:Q22992; EMBL:U64834; PIDN:AAB04823.1; GSPDB:GN00023; CESP:F54D11  
A;Experimental source: strain Bristol N2; clone F54D11  
C;Genetics:  
A;Gene: CESP:F54D11.2  
A;Map position: 5

A;Introns: 20/2; 279/3; 380/2; 755/3; 929/3; 1088/2

Query Match 51.4%; Score 55; DB 2; Length 1201;  
 Best Local Similarity 48.1%; Pred. No. 73;  
 Matches 13; Conservative 0; Mismatches 6; Indels 8; Gaps 1;

QY 2 PEKEEPTAAPAE-----RAPAPE 20  
 DB 187 PEKEEALAPPEPPAPQQAETPAPE 213

RESULT 13  
 A35085  
 trithorax protein - fruit fly (Drosophila melanogaster)  
 C;Species: Drosophila melanogaster  
 C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
 C;Accession: A35085; A38240  
 R;Mazo, A.M.; Huang, D.H.; Mozer, B.A.; Dawid, I.B.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 2112-2116, 1990  
 A;Title: The trithorax gene, a trans-acting regulator of the bithorax complex in Drosophila  
 A;Reference number: A35085; MUID:90192757; PMID:2107543  
 A;Accession: A35085  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-3759 <MAZ>  
 A;Cross-references: UNIPROT:P20659  
 R;Mazo, A.M.  
 submitted to GenBank, January 1990  
 A;Reference number: A38240  
 A;Accession: A38240  
 A;Molecule type: mRNA  
 A;Residues: 1-2361, 'Y', 2363-2397, 'N', 2399-2405, 'N', 2407-2411, 'N', 2413-3759 <MA2>  
 A;Cross-references: GB:M31617; NID:gi58817; PID:gi58818  
 C;Genetics:  
 A;Gene: FlyBase:trix  
 A;Cross-references: FlyBase:FBgn003862  
 C;Superfamily: histone methyltransferase, trithorax protein type  
 C;Keywords: DNA binding; transcription regulation; zinc finger

Query Match 51.4%; Score 55; DB 2; Length 3759;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 EEPTAAPAEAPAPE 20  
 DB 25 EDDAANPAEPQQAPE 40

RESULT 14  
 T44588  
 probable transcription regulator ty1Q [imported] - Streptomyces fradiae  
 C;Species: Streptomyces fradiae  
 C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
 C;Accession: T44588  
 R;Bate, N.; Butler, A.R.; Gandecha, A.R.; Cundliffe, E.  
 Chem. Biol. 6, 617-624, 1999  
 A;Title: Multiple regulatory genes in the tylosin-biosynthetic cluster of Streptomyces fradiae  
 A;Reference number: Z22801; MUID:99398833; PMID:10467127  
 A;Accession: T44588  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-213 <BAT>  
 A;Cross-references: UNIPROT:Q9XC05; EMBL:AF145049; PIDN:AAD40803.1  
 A;Experimental source: strain T59235  
 C;Genetics:  
 A;Gene: ty1Q

Query Match 50.5%; Score 54; DB 2; Length 213;  
 Best Local Similarity 55.8%; Pred. No. 18;  
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEAPAPE 19  
 DB 25 EDDAANPAEPQQAPE 40

Db 196 PRRARPGAAAGEPAAPAP 213

RESULT 15  
 H75253  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C;Species: Deinococcus radiodurans  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: H75253  
 R;White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: H75253  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-347 <WHI>  
 A;Cross-references: UNIPROT:Q9RR94; GB:AE002089; GB:AE000513; NID:g6460427; PIDN:AAF12144  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR2600  
 A;Map position: 1

Query Match 50.5%; Score 54; DB 2; Length 347;  
 Best Local Similarity 61.1%; Pred. No. 29;  
 Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 PEKEEPTAAPAEAPAPE 19  
 DB 46 PPSAPAAAPALPVAPAP 63

RESULT 16  
 B75398  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C;Species: Deinococcus radiodurans  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: B75398  
 R;White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: B75398  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-365 <WHI>  
 A;Cross-references: UNIPROT:Q9RUH9; GB:AE001986; GB:AE000513; NID:g6459162; PIDN:AAF10981  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR1407  
 A;Map position: 1

Query Match 50.5%; Score 54; DB 2; Length 365;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPEKEEPTAAPAEPEAPAPE 20  
 DB 336 TPPANQPAPAPQAPAPAPQ 355

RESULT 17  
 T08134  
 oleosin-like protein - rape  
 C;Species: Brassica napus (rape)  
 C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004  
 C;Accession: T08134  
 R;Hong, H.P.; Ross, J.H.; Gerster, J.L.; Rigas, S.; Datla, R.S.; Hatzopoulos, P.; Scoles,  
 Plant Mol. Biol. 34, 549-555, 1997



RESULT 20

S36152

MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - human

C/Species: Homo sapiens (man)

C/Date: 06-Jun-1995 #sequence\_revision 17-Nov-1995 #text\_change 15-Sep-2000

C/Accession: S36152

R/Iris, F.J.M.; Bougueleret, L.; Prieur, S.; Caterina, D.; Primas, G.; Perrot, V.; Jurka, Nature Genet. 3, 137-145, 1993

A/Title: Dense Alu clustering and a potential new member of the NFKappaB family within a

A/Reference number: S36152; MUID:93272029; PMID:8499947

A/Accession: S36152

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1872 <IRI>

A/Cross-references: EMBL:Z15025

A/Note: in the authors' translation residues 32-34 are shown after residue 4 and, consequ

C/Genetics:

A/Introns: 38/2; 97/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1; 651

C/Superfamily: collagen alpha 1(IV) chain

Query Match 50.5%; Score 54; DB 2; Length 1872;

Best Local Similarity 62.5%; Pred. No. 1.4e+02;

Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 KEEPTAAPAEPEAPAP 19

|||||

DB 505 KAEPAPAPAEPTAP 520

RESULT 21

B35098

MHC class III histocompatibility antigen HLA-B-associated protein 2 [imported] - human

C/Species: Homo sapiens (man)

C/Date: 10-Aug-1990 #sequence\_revision 06-Nov-1992 #text\_change 09-Jul-2004

C/Accession: B35098

R/Banerji, J.; Sands, J.; Strominger, J.L.; Spies, T. Proc. Natl. Acad. Sci. U.S.A. 87, 2374-2378, 1990

A/Title: A gene pair from the human major histocompatibility complex encodes large prolif

A/Reference number: A35098; MUID:90192810; PMID:2156268

A/Accession: B35098

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-2142 <BAN>

A/Cross-references: UNIPROT:P48634; GB:M33509; NID:g179338; PIDN:AAA35585.1; PID:g179333

A/Note: the authors translated the codon AGT for residue 97 as Gly

C/Superfamily: collagen alpha 1(IV) chain

Query Match 50.5%; Score 54; DB 2; Length 2142;

Best Local Similarity 62.5%; Pred. No. 1.6e+02;

Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 KEEPTAAPAEPEAPAP 19

|||||

DB 493 KAEPAPAPAEPTAP 508

RESULT 22

M0R4E

myosin alkali light chain 4, embryonic and atrial - rat

N/Alternate names: MLC1A; MLC1emb; myosin L1 catalytic light chain, atrial

C/Species: Rattus norvegicus (Norway rat)

C/Date: 29-Jan-1993 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004

C/Accession: S09236

R/Rovner, A.S.; McNally, E.M.; Leinwand, L.A. Nucleic Acids Res. 18, 1581-1586, 1990

A/Title: Complete cDNA sequence of rat atrial myosin light chain 1: patterns of expressi

A/Reference number: S09236; MUID:90221887; PMID:2326197

A/Accession: S09236

A/Molecule type: mRNA

A/Residues: 1-193 <ROV>





C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T32404  
 R:Le, T.T.; Kemp, K.; Scheet, P.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of C. elegans cosmid R148.  
 A:Reference number: Z21161  
 A:Accession: T32404  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1127 <LET>  
 A:Cross-references: UNIPROT:O17267; EMBL:AF025467; PIDN:AAB71038.1; GSPDB:GN00021; CESP:  
 A:Experimental source: strain Bristol N2; clone R148  
 C:Genetics:  
 A:Gene: CESP:R148.3  
 A:Map position: 3  
 A:Introns: 11/3; 42/1; 87/3; 373/3; 544/2; 595/2; 1042/1  
 C:Superfamily: Caenorhabditis elegans hypothetical protein R148.3

Query Match 49.5%; Score 53; DB 2; Length 1127;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 EPTAAPARPEAPAP 19  
 :|||:|:|:|:|:|:|  
 Db 294 QETTPAPLEAPAP 308

RESULT 33  
 T00519  
 prline-rich protein At2g2130 - Arabidopsis thaliana  
 N:Alternate names: protein T20D16.24  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T00519; H04620  
 R:Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, November 1997  
 A:Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
 A:Reference number: Z14164  
 A:Accession: T00519  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-185 <ROU>  
 A:Cross-references: UNIPROT:O22194; EMBL:AC002391; NID:G2642427; PID:G2642449  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
 Nature 402, 761-768, 1999  
 A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: H84620  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-185 <STO>  
 A:Cross-references: GB:AE002093; NID:G2642449; PIDN:AAB87117.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: T20D16.24; At2g23130  
 A:Map position: 2  
 A:Introns: 158/3

Query Match 48.6%; Score 52; DB 2; Length 185;  
 Best Local Similarity 45.0%; Pred. No. 26;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPKEEPTAAPARPEAPAP 20  
 :|||:|:|:|:|:|:|  
 Db 35 SPHKRPTSPAISPAAPTPE 54

RESULT 34  
 S29793  
 protein ORF 214 (atpA 3' region) - soybean mitochondrion

C:Species: mitochondrion Glycine max (soybean)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: S29793; S29794; S29874  
 R:Chanut, F.A.; Grabau, E.A.; Gesteland, R.F.  
 Curr. Genet. 23, 234-247, 1993  
 A:title: Complex organization of the soybean mitochondrial genome: recombination repeats  
 A:Reference number: S29792; MUID:93169687; PMID:8435853  
 A:Accession: S29793  
 A:Molecule type: DNA  
 A:Residues: 1-214 <CHAL>  
 A:Cross-references: UNIPROT:Q01925; EMBL:Z14031; NID:G22738; PIDN:CAA78406.1; PID:G22740  
 A:Accession: S29794  
 A:Molecule type: protein  
 A:Residues: 109-202 <CHA2>  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: mitochondrion

Query Match 48.6%; Score 52; DB 2; Length 214;  
 Best Local Similarity 55.0%; Pred. No. 30;  
 Matches 11; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY 1 TPKEEPTAAPARPEAPAP 20  
 :|||:|:|:|:|:|:|  
 Db 60 TPD---PAAPAEPOAPDPD 75

RESULT 35  
 T35351  
 hypothetical protein SC5H1.35c - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
 C:Accession: T35351  
 R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, May 1999  
 A:Reference number: Z21575  
 A:Accession: T35351  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-237 <OLI>  
 A:Cross-references: UNIPROT:Q9X7U2; EMBL:AL049863; PIDN:CAB42960.1; GSPDB:GN00070; SCOEDE  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDEB:SC5H1.35c

Query Match 48.6%; Score 52; DB 2; Length 237;  
 Best Local Similarity 66.7%; Pred. No. 34;  
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 EPTAAPARPEAPAP 20  
 :|||:|:|:|:|:|:|  
 Db 167 EPVRPPAPEPSAPE 181

RESULT 36  
 H71563  
 hypothetical protein CT050 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
 C:Accession: H71563  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
 Science 282, 754-759, 1998  
 A:title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tract  
 A:Reference number: A71570; MUID:99000809; PMID:9784136  
 A:Accession: H71563  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-536 <ARN>  
 A:Cross-references: UNIPROT:O84053; GB:AE001279; GB:AE001273; NID:G3328434; PIDN:AAC67641  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: CT050

```
Query Match      48.6%; Score 52; DB 2; Length 536;
Best Local Similarity 64.3%; Pred. No. 74;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTAAPAEPEAPE 20
DB 397 PESPPSPPEAPE 410
      | : | | | | |
      | : | | | | |

RESULT 37
T43556
Wiskott-Aldrich syndrome protein homolog - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43556
R;Zankel, T.C.; Ow, D.W.
Submitted to the EMBL Data Library, December 1997
A;Description: A Wiskott-Aldrich Syndrome protein homolog in Schizosaccharomyces pombe.
A;Reference number: Z22575
A;Accession: T43556
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-574 <ZAN>
A;Cross-references: UNIPROT:O36027; EMBL:AF038575; PIDN:AAB92587.1
A;Experimental source: strain JS21
C;Genetics:
A;Gene: wapl
A;Map position: 1
A;Introns: 72/3; 519/3; 564/1

Query Match      48.6%; Score 52; DB 2; Length 574;
Best Local Similarity 76.9%; Pred. No. 79;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTAAPAEPEAPAP 19
DB 470 PPAAPAPPPAPAP 482
      | | | | | | |
      | | | | | | |

RESULT 38
T38819
wiskott-aldrich syndrome protein homolog 1 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38819
R;Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
Submitted to the EMBL Data Library, September 1997
A;Reference number: Z21813
A;Accession: T38819
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-574 <CON>
A;Cross-references: UNIPROT:O36027; EMBL:Z96980; NID:e1060691; PIDN:CAB11718.1; GSPDB:GN
A;Experimental source: strain 972h-; cosmid c4F10
C;Genetics:
A;Gene: wapl; SPDB:SPAC4F10.15C
A;Map position: 1
A;Introns: 72/3; 519/3; 564/1

Query Match      48.6%; Score 52; DB 2; Length 574;
Best Local Similarity 76.9%; Pred. No. 79;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTAAPAEPEAPAP 19
DB 470 PPAAPAPPPAPAP 482
      | | | | | | |
      | | | | | | |

RESULT 39
G83146
penicillin-binding protein 2 PA4003 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
```

```
C;Accession: G83146
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
C;Accession: G83146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-646 <STO>
A;Cross-references: UNIPROT:Q9X6V3; GB:AE004817; GB:AE004091; NID:g9950181; PIDN:AAG07390
A;Experimental source: strain PA01
C;Genetics:
A;Gene: pb; PA4003
C;Superfamily: penicillin-binding protein 3

Query Match      48.6%; Score 52; DB 2; Length 646;
Best Local Similarity 76.5%; Pred. No. 88;
Matches 13; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 4 KEEPTAAPAEPEAPE 20
DB 631 KPEPTA-AEPEAPE 645
      | | | | | | |
      | | | | | | |

RESULT 40
T43952
hypothetical protein DR2 [imported] - human herpesvirus 6
C;Species: human herpesvirus 6
A;Variety: strain HST
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43952; T44059
R;Isegawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawar
J. Virol. 73, 8053-8063, 1999
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and
A;Reference number: Z22732; MUID:99412319; PMID:10482554
A;Accession: T43952
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-647 <ISE>
A;Cross-references: UNIPROT:Q9W9G4; EMBL:AB021506; NID:g4995977; PIDN:BAA78213.1; PID:G
A;Experimental source: strain HST; pop. variant B
A;Accession: T44059
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-647 <IS2>
A;Cross-references: EMBL:AB021506; NID:g4995977; PIDN:BAA78213.1; PID:g4995980
A;Experimental source: strain HST; pop. variant B
C;Genetics:
A;Gene: DR2; DR2R

Query Match      48.6%; Score 52; DB 2; Length 647;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20
DB 488 TPVEDEFSARPCPPGPAEE 507
      | | | | | | |
      | | | | | | |

Search completed: August 11, 2005, 09:58:30
Job time : 41 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: August 11, 2005, 09:49:16 ; Search time 43 Seconds  
(without alignments)  
34.720 Million cell updates/sec

Title: US-10-089-273-7

Perfect score: 107

Sequence: 1 TPEKEPTAAPAEPAPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 63.5  | 59.3        | 157    | 4     | US-09-902-540-13049  |
| 2          | 61    | 57.0        | 666    | 4     | US-09-050-739-70     |
| 3          | 59    | 55.1        | 174    | 4     | US-09-248-796A-25211 |
| 4          | 56.5  | 52.8        | 356    | 4     | US-09-902-540-16059  |
| 5          | 56.5  | 52.8        | 605    | 4     | US-08-714-741-46     |
| 6          | 56    | 52.3        | 181    | 4     | US-08-529-055-42     |
| 7          | 56    | 52.3        | 605    | 2     | US-08-687-956A-1     |
| 8          | 56    | 52.3        | 8991   | 4     | US-08-714-741-32     |
| 9          | 55.5  | 51.9        | 197    | 4     | US-08-529-055-44     |
| 10         | 55.5  | 51.9        | 344    | 4     | US-09-902-540-12463  |
| 11         | 55    | 51.4        | 168    | 4     | US-08-529-055-55     |
| 12         | 55    | 51.4        | 183    | 4     | US-08-529-055-50     |
| 13         | 55    | 51.4        | 941    | 4     | US-07-757-022B-14    |
| 14         | 55    | 51.4        | 1022   | 4     | US-07-757-022B-84    |
| 15         | 55    | 51.4        | 1038   | 4     | US-07-757-022B-74    |
| 16         | 55    | 51.4        | 1049   | 4     | US-07-757-022B-58    |
| 17         | 55    | 51.4        | 1140   | 4     | US-07-757-022B-104   |
| 18         | 55    | 51.4        | 1270   | 4     | US-07-757-022B-44    |
| 19         | 55    | 51.4        | 1311   | 4     | US-07-757-022B-42    |
| 20         | 55    | 51.4        | 1313   | 4     | US-07-757-022B-142   |
| 21         | 55    | 51.4        | 1314   | 4     | US-07-757-022B-50    |
| 22         | 55    | 51.4        | 1320   | 4     | US-07-757-022B-46    |
| 23         | 55    | 51.4        | 1320   | 4     | US-07-757-022B-60    |
| 24         | 55    | 51.4        | 1320   | 4     | US-10-164-595-58     |
| 25         | 55    | 51.4        | 1354   | 4     | US-07-757-022B-48    |
| 26         | 55    | 51.4        | 1361   | 4     | US-07-757-022B-40    |
| 27         | 55    | 51.4        | 1363   | 4     | US-07-757-022B-52    |

28 55 51.4 1404 4 US-07-757-022B-2 Sequence 2, Appli  
29 55 51.4 1404 4 US-07-757-022B-62 Sequence 62, Appli  
30 55 51.4 1404 4 US-10-164-595-78 Sequence 78, Appli  
31 55 51.4 1404 4 US-09-298-970A-1 Sequence 1, Appli  
32 55 51.4 1411 4 US-09-949-016-10827 Sequence 10827, A  
33 54.5 50.9 213 4 US-08-529-055-47 Sequence 47, Appli  
34 54 50.5 2142 4 US-09-538-092-1142 Sequence 1142, Ap  
35 53.5 50.0 185 4 US-08-529-055-67 Sequence 67, Appli  
36 53.5 50.0 211 4 US-08-529-055-68 Sequence 68, Appli  
37 53.5 50.0 212 4 US-08-529-055-68 Sequence 10376, A  
38 53 49.5 211 4 US-09-902-540-10376 Sequence 32760, A  
39 53 49.5 551 4 US-09-252-991A-32760 Sequence 6777, Ap  
40 53 49.5 1026 4 US-09-949-016-6777 Sequence 10870, A  
41 53 49.5 1034 4 US-09-949-016-10870 Sequence 4, Appli  
42 53 49.5 1219 4 US-09-344-624-4 Sequence 59, Appli  
43 52.5 49.1 170 4 US-08-529-055-59 Sequence 58, Appli  
44 52.5 49.1 188 4 US-08-529-055-58  
45 52.5 49.1 204 4 US-08-529-055-58

## ALIGNMENTS

RESULT 1  
US-09-902-540-13049  
; Sequence 13049, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkie, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR FILING NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 13049  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
; US-09-902-540-13049

Query Match 59.3% Score 63.5; DB 4; Length 157;  
Best Local Similarity 70.0%; Pred No. 1.6;  
Matches 14; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 2 PEKEPTAAPAEPAPE 20  
||: |||||  
Db 122 PEEAEASAAATPEAPVPE 141

RESULT 2  
US-09-050-739-70  
; Sequence 70, Application US/09050739  
; Patent No. 6641814  
; GENERAL INFORMATION:  
; APPLICANT: ANDERSEN, Peter  
; APPLICANT: NIELSEN, Rikke  
; APPLICANT: OETTINGER, Thomas  
; APPLICANT: RASMUSSEN, Peter Birk  
; APPLICANT: ROSENKRANDS, Ida  
; APPLICANT: WELDLINGH, Karin  
; APPLICANT: FLORIO, Walter  
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS  
; FILE REFERENCE: 670001-2002.1  
; CURRENT APPLICATION NUMBER: US/09/050,739  
; CURRENT FILING DATE: 1998-03-30  
; EARLIER APPLICATION NUMBER: 0376/97  
; EARLIER FILING DATE: 1997-04-02

```
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-70

Query Match      57.0%; Score 61; DB 4; Length 666;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEPAPE 19
Db 156 FTPPTGTGAPQPESEAP 173

RESULT 3
US-09-248-796A-25211
; Sequence 25211, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25211
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-25211

Query Match      55.1%; Score 59; DB 4; Length 174;
Best Local Similarity 57.9%; Pred. No. 6;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEPAPE 20
Db 68 PEQPEPSAPAPASAPAE 86

RESULT 4
US-09-902-540-16059
; Sequence 16059, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barty S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16059
; LENGTH: 356
; TYPE: PRT
```

```
; ORGANISM: Myxococcus xanthus
US-09-902-540-16059

Query Match      52.8%; Score 56.5; DB 4; Length 356;
Best Local Similarity 63.2%; Pred. No. 25;
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 PEKEPTAAPAEPAPE 20
Db 267 PPEEPAPPPPEP-APAE 284

RESULT 5
US-08-714-741-46
; Sequence 46, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF.
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-46

Query Match      52.8%; Score 56.5; DB 4; Length 605;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 EKEEPTAAPAEPAPE 20
Db 524 EKEKPAPAEKP-APAE 540

RESULT 6
US-08-529-055-42
; Sequence 42, Application US/08529055
```



Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 181 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-42

Query Match 52.3%; Score 56; DB 4; Length 181;  
Best Local Similarity 52.0%; Pred. No. 14;  
Matches 13; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

Qy 2 PEK-----EPTAAPAEPEAPE 20  
Db 131 PEKPAEKPAEKPAEKPAEKPAPE 155

RESULT 7  
US-08-687-956A-1  
Sequence 1, Application US/08687956A  
Patent No. 5861157  
GENERAL INFORMATION:  
APPLICANT: BURNIE, JAMES P  
APPLICANT: MATTHEWS, RUTH C  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
TITLE OF INVENTION: INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP  
STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH  
STREET: FLOOR  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,956A  
FILING DATE: 29-JUL-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9401689.6  
FILING DATE: 28-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 50885/222892  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202/861-3000  
TELEFAX: 202/822-0944  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 605 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus oralis  
US-08-687-956A-1

Query Match 52.3%; Score 56; DB 2; Length 605;  
Best Local Similarity 55.0%; Pred. No. 49;  
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TPEKEPTAAPAEPEAPE 20  
Db 158 TKEKTPPPVAPTEPQAPTE 177

RESULT 8  
US-08-714-741-32  
Sequence 32, Application US/08714741  
Patent No. 6500613  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J.  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
TITLE OF INVENTION: PORTIONS AND PRODUCTS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,741  
FILING DATE: 16-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 168 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-55

Query Match 51.4%; Score 55; DB 4; Length 168;  
Best Local Similarity 55.0%; Pred. No. 17;  
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TPEKEEPTAAPEAPEAPE 20  
Db 115 TPEAPAPAPAPKAPAPK 134

## RESULT 12

US-08-529-055-50  
Sequence 50, Application US/08529055  
Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Therof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 183 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-50

Query Match 51.4%; Score 55; DB 4; Length 183;  
Best Local Similarity 55.0%; Pred. No. 18;  
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TPEKEEPTAAPEAPEAPE 20

Db 115 TPEAPAPAPAPKAPAPK 134

## RESULT 13

US-07-757-022B-14  
Sequence 14, Application US/07757022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.  
APPLICANT: Clark, Stephen C.  
APPLICANT: Turner, Katherine  
APPLICANT: Hewick, Rodney M.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
PRIOR APPLICATION DATA: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 941 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-757-022B-14

Query Match 51.4%; Score 55; DB 4; Length 941;  
Best Local Similarity 52.2%; Pred. No. 1e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

Qy 1 TPEKEEPTA---APAEPEAPAP 19  
Db 437 TPEKPAPTPEELAPTTPEEPTP 459

## RESULT 14

US-07-757-022B-84  
Sequence 84, Application US/07757022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.

```
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 18-JAN-1991
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-84

Query Match 51.4%; Score 55; DB 4; Length 1022;
Best Local Similarity 52.2%; Pred. No. 1.le+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 1 TPEKEPTA----APAEPEAPAP 19
Db 518 TPEKPATTPPELAPTPEPTP 540

RESULT 15
US-07-757-022B-74
; Sequence 74, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
```

```
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1038 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-74

Query Match 51.4%; Score 55; DB 4; Length 1038;
Best Local Similarity 52.2%; Pred. No. 1.le+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 1 TPEKEPTA----APAEPEAPAP 19
Db 502 TPEKPATTPPELAPTPEPTP 524

RESULT 16
US-07-757-022B-58
; Sequence 58, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
```

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA: US 07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
PRIOR APPLICATION DATA: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1049 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-757-022B-58

Query Match 51.4%; Score 55; DB 4; Length 1049;  
Best Local Similarity 52.2%; Pred. No. 1.1e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

Qy 1 TPEKEPTA---APAEPEAPAP 19  
Db 545 TPEKPATTPEELAPTTPEPTP 567

RESULT 17  
US-07-757-022B-104  
Sequence 104, Application US/07/57022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.  
APPLICANT: Clark, Stephen C.  
APPLICANT: Turner, Katherine  
APPLICANT: Hewick, Rodney M.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1140 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-757-022B-104

Query Match 51.4%; Score 55; DB 4; Length 1140;  
Best Local Similarity 52.2%; Pred. No. 1.2e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

Qy 1 TPEKEPTA---APAEPEAPAP 19  
Db 636 TPEKPATTPEELAPTTPEPTP 658

RESULT 18  
US-07-757-022B-44  
Sequence 44, Application US/07/57022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.  
APPLICANT: Clark, Stephen C.  
APPLICANT: Turner, Katherine  
APPLICANT: Hewick, Rodney M.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851

```

; INFORMATION FOR SEQ ID NO: 44:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-757-022B-44

```

Query Match 51.4%; Score 55; DB 4; Length 1270;  
Best Local Similarity 52.2%;  
Pred. NO. 1.4e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels

QY 1 TPEKEPTA---APAEPEAP 19  
||| ||| ||| ||| |||  
Db 502 TPEKPATTPPEELAPTTPEEPTP 524

```

RESULT 19
US-07-757-022B-42
; Sequence 42, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.

```

|                       |        |                    |       |              |
|-----------------------|--------|--------------------|-------|--------------|
| Query Match           | 51.4%; | Score 55;          | DB 4; | Length 1311; |
| Best Local Similarity | 52.2%; | Pred. No. 1.4e+02; |       |              |

|    | Matches | 12; Conservative         | 0; Mismatches | 7; Indels | 4; Gaps |
|----|---------|--------------------------|---------------|-----------|---------|
| QY | 1       | TPEKGEPTA-----APAEPEAPAP | 19            |           |         |
|    |         |                          |               |           |         |
| Db | 543     | TPEKAPTTPPEELAPTTPEEPTP  | 565           |           |         |

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

```

Query Match      51.4%; Score 55; DB 4; Length 1313;
Best Local Similarity 52.2%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 1 TPEKEEPTA----APAEPEAPAP 19
      |||||
Db 545 TPEKPAPTTPEELAPTTPEEPT 567

```

RESULT 21  
US-07-757-022B-50  
: Sequence 50, Application US/07757022B

```

RESULT 23
US-07-757-022B-60
; Sequence 60, Application US/07757022B
; Patent No. 6431142
; GENERAL INFORMATION:
; APPLICANT: Geeser, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/757,022B  
; FILING DATE: 19910910  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1320 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-757-022B-60

Query Match 51.4%; Score 55; DB 4; Length 1320;  
Best Local Similarity 52.2%; Pred. No. 1.4e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;  
  
QY 1 TPEKPEPTA----APAEPEAPAP 19  
|||||  
Db 552 TPEKPAPTPEELAPTTPEPTP 574

RESULT 24  
US-10-164-595-58  
; Sequence 58, Application US/10164595  
; Patent No. 6657054  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc  
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides  
; FILE REFERENCE: 1U 103 R1  
; CURRENT APPLICATION NUMBER: US/10/164,595  
; CURRENT FILING DATE: 2002-06-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58  
; LENGTH: 1320  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-164-595-58

Query Match 51.4%; Score 55; DB 4; Length 1320;  
Best Local Similarity 52.2%; Pred. No. 1.4e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;  
  
QY 1 TPEKPEPTA----APAEPEAPAP 19  
|||||  
Db 552 TPEKPAPTPEELAPTTPEPTP 574

RESULT 25  
US-07-757-022B-48  
; Sequence 48, Application US/07757022B  
; Patent No. 6433142  
; GENERAL INFORMATION:  
; APPLICANT: Gesner, Thomas G.

;  
; APPLICANT: Clark, Stephen C.  
; APPLICANT: Turner, Katherine  
; APPLICANT: Hewick, Rodney M.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/757,022B  
; FILING DATE: 19910910  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1354 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-757-022B-48  
  
Query Match 51.4%; Score 55; DB 4; Length 1354;  
Best Local Similarity 52.2%; Pred. No. 1.5e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;  
  
QY 1 TPEKPEPTA----APAEPEAPAP 19  
|||||  
Db 586 TPEKPAPTPEELAPTTPEPTP 608  
  
RESULT 26  
US-07-757-022B-40  
; Sequence 40, Application US/07757022B  
; Patent No. 6433142  
; GENERAL INFORMATION:  
; APPLICANT: Gesner, Thomas G.  
; APPLICANT: Clark, Stephen C.  
; APPLICANT: Turner, Katherine  
; APPLICANT: Hewick, Rodney M.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts



COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER: IBM PC compatible  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US/07/390,901  
FILING DATE: 08-AUG-1989  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1361 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-757-022B-40

Query Match 51.4%; Score 55; DB 4; Length 1361;  
Best Local Similarity 52.2%; Pred. No. 1.5e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 1 TPEKEPTA---APAEPEAPAP 19  
Db 593 TPEKPATTPEELAPTTPEPTP 615

RESULT 27  
US-07-757-022B-52  
Sequence 52, Application US/07757022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.  
APPLICANT: Clark, Stephen C.  
APPLICANT: Turner, Katherine  
APPLICANT: Hewick, Rodney M.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US/07/390,901  
FILING DATE: 08-AUG-1989  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1363 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-757-022B-52

Query Match 51.4%; Score 55; DB 4; Length 1363;  
Best Local Similarity 52.2%; Pred. No. 1.5e+02;  
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 1 TPEKEPTA---APAEPEAPAP 19  
Db 595 TPEKPATTPEELAPTTPEPTP 617

RESULT 28  
US-07-757-022B-2  
Sequence 2, Application US/07757022B  
Patent No. 6433142  
GENERAL INFORMATION:  
APPLICANT: Gesner, Thomas G.  
APPLICANT: Clark, Stephen C.  
APPLICANT: Turner, Katherine  
APPLICANT: Hewick, Rodney M.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,022B  
FILING DATE: 19910910  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/643,502  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/546,114  
FILING DATE: 29-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/457,196  
FILING DATE: 29-DEC-1989

```

; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-757-022B-62

Query Match          51.4%; Score 55; DB 4; Length 1404;
Best Local Similarity 52.2%; Pred. No. 1.5e+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY      1 TPEKEPTA----APAEPEAPAP 19
          |||| || || || || || ||
Db      636 TPEKPAPTTPEELAPTTPEPTP 658

RESULT 30
US-10-164-595-78
; Sequence 78, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: IU 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-78

Query Match          51.4%; Score 55; DB 4; Length 1404;
Best Local Similarity 52.2%; Pred. No. 1.5e+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY      1 TPEKEPTA----APAEPEAPAP 19
          |||| || || || || || ||
Db      636 TPEKPAPTTPEELAPTTPEPTP 658

RESULT 31
US-09-298-970A-1
; Sequence 1, Application US/09298970A
; Patent No. 6743774
; GENERAL INFORMATION:
; APPLICANT: Jav, Gregory D.
; TITLE OF INVENTION: TRIBONECTINS
; FILE REFERENCE: 21486-026
; CURRENT APPLICATION NUMBER: US/09/298,970A
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-970A-1

Query Match          51.4%; Score 55; DB 4; Length 1404;
Best Local Similarity 52.2%; Pred. No. 1.5e+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY      1 TPEKEPTA----APAEPEAPAP 19
          |||| || || || || || ||
Db      636 TPEKPAPTTPEELAPTTPEPTP 658

RESULT 32
US-09-949-016-10827

```

```
; Sequence 10827, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10827
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10827

Query Match 51.4%; Score 55; DB 4; Length 1411;
Best Local Similarity 52.2%; Pred. No. 1.5e+02;
Matches 12; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

Qy 1 TPEKEPTA-----APAEPEAPAP 19
Db 643 TPEKPAPTPEELAPTTPEEFTP 665

RESULT 33
US-08-529-055-47
; Sequence 47, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-47

Query Match 50.9%; Score 54.5; DB 4; Length 213;
Best Local Similarity 57.9%; Pred. No. 25;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 2 PEKEPTAAPA-EPEAPAP 19
Db 170 PKPEQPAPAPAPKEQPAP 189

RESULT 34
US-09-538-092-1142
; Sequence 1142, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1142
; LENGTH: 2142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P48634
; US-09-538-092-1142

Query Match 50.5%; Score 54; DB 4; Length 2142;
Best Local Similarity 62.5%; Pred. No. 3.1e+02;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 KEPEPTAAPPEAPAP 19
Db 493 KAEPAAPPAAPSTPAP 508

RESULT 35
US-08-529-055-69
; Sequence 69, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-69

Query Match 50.0%; Score 53.5; DB 4; Length 185;
Best Local Similarity 57.9%; Pred. No. 28;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPA-EPEAPAP 19
Db 164 PRAEKPAPAPKPEQPAP 182

RESULT 36
US-08-529-055-67
; Sequence 67, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-68

Query Match 50.0%; Score 53.5; DB 4; Length 212;
Best Local Similarity 57.9%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPA-EPEAPAP 19
Db 191 PRAEKPAPAPKPEQPAP 209

RESULT 37
US-08-529-055-68
; Sequence 68, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-68

Query Match 50.0%; Score 53.5; DB 4; Length 212;
Best Local Similarity 57.9%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPA-EPEAPAP 19
Db 191 PRAEKPAPAPKPEQPAP 209
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 211 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-67

Query Match 50.0%; Score 53.5; DB 4; Length 211;
Best Local Similarity 57.9%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPA-EPEAPAP 19
Db 188 PRAEKPAPAPKPEQPAP 206

RESULT 37
US-08-529-055-68
; Sequence 68, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-68

Query Match 50.0%; Score 53.5; DB 4; Length 212;
Best Local Similarity 57.9%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 PEKEEPTAAPA-EPEAPAP 19
Db 191 PRAEKPAPAPKPEQPAP 209
```

```
RESULT 38
US-09-902-540-10376
; Sequence 10376, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10376
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10376

Query Match          49.5%; Score 53; DB 4; Length 211;
Best Local Similarity 52.6%; Pred. No. 36;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      1 TPEKEEPTAAPAEAPAP 19
      | : ||| ||| |||
Db      190 TLDSEVPSAAPAPAGPKP 208

RESULT 39
US-09-252-991A-32760
; Sequence 32760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32760
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32760

Query Match          49.5%; Score 53; DB 4; Length 551;
Best Local Similarity 58.8%; Pred. No. 98;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      4 KEEPTAAPAEAPAE 20
      | : ||| ||| |||
Db      206 EEQPAAPAAQAAPAAE 222

RESULT 40
US-09-949-016-6777
; Sequence 6777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6777
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6777

Query Match          49.5%; Score 53; DB 4; Length 1026;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 TPEKEEPTAAPAEAPAE 20
      | : ||| ||| |||
Db      33 TPESTTEVKAPPAPPPPPPE 52

Search completed: August 11, 2005, 09:59:18
Job time : 44 secs
```

This page Blank (uspto)

Nolan.P.  
10/089273  
Seq.107

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2005, 09:38:55 ; Search time 163 Seconds  
(without alignments)  
47.455 Million cell updates/sec

Title: US-10-089-273-7

Perfect score: 107

Sequence: 1 TPEKEBPTRAPAEPEAPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 107   | 100.0       | 20     | AAU01421  | AAU01421 Peptide #  |
| 2          | 107   | 100.0       | 151    | AAU01415  | AAU01415 Latex all  |
| 3          | 102   | 95.3        | 19     | AAU01444  | AAU01444 Peptide #  |
| 4          | 102   | 95.3        | 19     | AAU01432  | AAU01432 Peptide #  |
| 5          | 95    | 88.8        | 18     | AAU01433  | AAU01433 Peptide #  |
| 6          | 95    | 88.8        | 18     | AAU01445  | AAU01445 Peptide #  |
| 7          | 91    | 85.0        | 17     | AAU01434  | AAU01434 Peptide #  |
| 8          | 90    | 84.1        | 17     | AAU01446  | AAU01446 Peptide #  |
| 9          | 85    | 79.4        | 16     | AAU01447  | AAU01447 Peptide #  |
| 10         | 84    | 78.5        | 16     | AAU01435  | AAU01435 Peptide #  |
| 11         | 80    | 74.8        | 15     | AAU01448  | AAU01448 Peptide #  |
| 12         | 80    | 74.8        | 15     | AAU01436  | AAU01436 Peptide #  |
| 13         | 75    | 70.1        | 14     | AAU01437  | AAU01437 Peptide #  |
| 14         | 75    | 70.1        | 14     | AAU01449  | AAU01449 Peptide #  |
| 15         | 68    | 63.6        | 13     | AAU01438  | AAU01438 Peptide #  |
| 16         | 68    | 63.6        | 13     | AAU01450  | AAU01450 Peptide #  |
| 17         | 63    | 58.9        | 12     | AAU01439  | AAU01439 Peptide #  |
| 18         | 63    | 58.9        | 12     | AAU01451  | AAU01451 Peptide #  |
| 19         | 63    | 58.9        | 4412   | AAV53666  | AAV53666 Sequence   |
| 20         | 63    | 58.9        | 34350  | ADQ89964  | ADQ89964 Antagonis  |
| 21         | 61    | 57.0        | 666    | AAW72911  | AAW72911 Mycobacte  |
| 22         | 61    | 57.0        | 666    | AAV21928  | AAV21928 Amino aci  |
| 23         | 61    | 57.0        | 666    | AAAB19846 | AAAB19846 Mycobacte |
| 24         | 61    | 57.0        | 666    | ABU05989  | ABU05989 M. tuberc  |
| 25         | 61    | 57.0        | 666    | ABU37020  | ABU37020 Protein e  |

|    |      |      |      |   |          |                     |
|----|------|------|------|---|----------|---------------------|
| 26 | 59.5 | 55.6 | 1585 | 7 | ABM85412 | ABM85412 Mouse pro  |
| 27 | 59.5 | 55.6 | 2462 | 6 | ABU61813 | ABU61813 Mouse nuc  |
| 28 | 59   | 55.1 | 11   | 4 | AAU01452 | AAU01452 Peptide #  |
| 29 | 59   | 55.1 | 11   | 4 | AAU01440 | AAU01440 Peptide #  |
| 30 | 59   | 55.1 | 20   | 4 | AAU01420 | AAU01420 Peptide #  |
| 31 | 59   | 55.1 | 20   | 4 | AAU01422 | AAU01422 Peptide #  |
| 32 | 59   | 55.1 | 20   | 5 | ABB84377 | ABB84377 H. brasili |
| 33 | 59   | 55.1 | 382  | 5 | ABG97431 | ABG97431 S. ghanae  |
| 34 | 59   | 55.1 | 428  | 6 | ABU34118 | ABU34118 Protein e  |
| 35 | 57   | 53.3 | 267  | 5 | ABG60119 | ABG60119 Human DIT  |
| 36 | 57   | 53.3 | 379  | 5 | ABP25792 | ABP25792 Streptoco  |
| 37 | 57   | 53.3 | 379  | 6 | ABU46857 | ABU46857 Protein e  |
| 38 | 57   | 53.3 | 379  | 8 | ADR83966 | ADR83966 S. pyogen  |
| 39 | 57   | 53.3 | 497  | 7 | ADC72877 | ADC72877 HIV-1 gag  |
| 40 | 57   | 53.3 | 746  | 5 | ADI28089 | ADI28089 Human nuc  |
| 41 | 57   | 53.3 | 764  | 5 | AAO17706 | AAO17706 Human PPA  |
| 42 | 57   | 53.3 | 959  | 8 | ADQ65779 | ADQ65779 Novel hum  |
| 43 | 57   | 53.3 | 972  | 7 | ADC42842 | ADC42842 RENAP pro  |
| 44 | 57   | 53.3 | 996  | 5 | AAE22760 | AAE22760 Human PGC  |
| 45 | 57   | 53.3 | 1002 | 5 | AAE22759 | AAE22759 Human PGC  |

## ALIGNMENTS

RESULT 1  
AAU01421  
ID AAU01421 standard; peptide; 20 AA.

AC AAU01421;

DT 18-JUL-2001 (first entry)

XX Peptide #7 derived from natural latex rubber allergen Hev b 5.

DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;

XX latex hypersensitivity.

KW Hevea brasiliensis.

OS WO200123429-A1.

XX 05-APR-2001.

XX 22-SEP-2000; 2000WO-AU001182.

XX 24-SEP-1999; 99AU-00003057.

PR 24-JUL-2000; 2000AU-00008964.

XX (MONU ) UNIV MONASH.

PA (REAS-) COOP RES CENT ASTHMA.

XX O'hehir R;

XX WPI; 2001-258129/26.

XX New peptide and immunospecific antibodies, useful for treatment and/or

PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence

PT derived from or homologous to Hev b 5.

XX Claim 8; Fig 1; 101pp; English.

XX The sequence represents the amino acid sequence of peptide #7 derived

CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.

CC The peptide is capable of interacting with T cells and modifying T cell

CC function when incubated with cells from subjects with a condition

CC characterised by an aberrant, unwanted or inappropriate immune response

CC to Hev b 5. The peptides, or antibodies to these peptides are useful for

CC the manufacture of a medicament for the treatment and/or prophylaxis of

CC latex hypersensitivity in a subject. They are also useful for diagnosing

CC or monitoring latex hypersensitivity in a subject

XX Sequence 20 AA;

SQ

```

Query Match      100.0%; Score 107; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20
   |||||
Db 1 TPEKEPTAAPAEPEAPE 20
   |||||

RESULT 2
AAU01415
ID AAU01415 standard; protein; 151 AA.
XX
AC AAU01415;
XX
DT 18-JUL-2001 (first entry)
XX
DE Latex allergen Hev b 5 amino acid sequence.
XX
KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;
KW latex hypersensitivity.
XX
OS Hevea brasiliensis.
XX
PN WO200123429-A1.
XX
PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000WO-AU001182.
XX
PR 24-SEP-1999; 99AU-00003057.
PR 24-JUL-2000; 2000AU-00008964.
XX
PA (MONU ) UNIV MONASH.
PA (REAS-) COOP RES CENT ASTHMA.
XX
PI O'hehir R;
XX
DR WPI; 2001-258129/26.
XX
PT New peptide and immunospecific antibodies, useful for treatment and/or
PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence
PT derived from or homologous to Hev b 5.
XX
PS Example 1; Page 89-90; 101pp; English.
XX
CC The sequence represents the amino acid sequence of natural rubber latex
CC allergen Hev b 5, isolated from rubber plant, Hevea brasiliensis. The
CC sequence was used to design peptides capable of interacting with T cells
CC and modifying T cell function when incubated with cells from subjects
CC with a condition characterised by an aberrant, unwanted or inappropriate
CC immune response to Hev b 5. The peptides, or antibodies to these peptides
CC are useful for the manufacture of a medicament for the treatment and/or
CC prophylaxis of latex hypersensitivity in a subject. They are also useful
CC for diagnosing or monitoring latex hypersensitivity in a subject
XX
SQ Sequence 151 AA;

Query Match      100.0%; Score 107; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20
   |||||
Db 46 TPEKEPTAAPAEPEAPE 65
   |||||

RESULT 3
AAU01444
ID AAU01444 standard; peptide; 19 AA.
XX
AC AAU01444;

```

```

XX
DT 18-JUL-2001 (first entry)
XX
DE Peptide #30 derived from natural latex rubber allergen Hev b 5.
XX
KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;
KW latex hypersensitivity.
XX
OS Hevea brasiliensis.
XX
PN WO200123429-A1.
XX
PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000WO-AU001182.
XX
PR 24-SEP-1999; 99AU-00003057.
PR 24-JUL-2000; 2000AU-00008964.
XX
PA (MONU ) UNIV MONASH.
PA (REAS-) COOP RES CENT ASTHMA.
XX
PI O'hehir R;
XX
DR WPI; 2001-258129/26.
XX
PT New peptide and immunospecific antibodies, useful for treatment and/or
PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence
PT derived from or homologous to Hev b 5.
XX
PS Disclosure; Fig 10; 101pp; English.
XX
CC The sequence represents the amino acid sequence of peptide #30 derived
CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.
CC The peptide is capable of interacting with T cells and modifying T cell
CC function when incubated with cells from subjects with a condition
CC characterised by an aberrant, unwanted or inappropriate immune response
CC to Hev b 5. The peptides, or antibodies to these peptides are useful for
CC the manufacture of a medicament for the treatment and/or prophylaxis of
CC latex hypersensitivity in a subject. They are also useful for diagnosing
CC or monitoring latex hypersensitivity in a subject
XX
SQ Sequence 19 AA;

Query Match      95.3%; Score 102; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PEKEPTAAPAEPEAPE 20
   |||||
Db 1 PEKEPTAAPAEPEAPE 19
   |||||

RESULT 4
AAU01432
ID AAU01432 standard; peptide; 19 AA.
XX
AC AAU01432;
XX
DT 18-JUL-2001 (first entry)
XX
DE Peptide #18 derived from natural latex rubber allergen Hev b 5.
XX
KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;
KW latex hypersensitivity.
XX
OS Hevea brasiliensis.
XX
PN WO200123429-A1.
XX
PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000WO-AU001182.

```



XX 24-SEP-1999; 99AU-00003057.  
 PR 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 DR New peptide and immunospecific antibodies, useful for treatment and/or  
 XX prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 PT Disclosure; Fig 10; 101pp; English.  
 PS The sequence represents the amino acid sequence of peptide #18 derived  
 XX from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 19 AA;  
 SQ

Query Match 95.3%; Score 102; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. NO. 3.1e-05;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAPAPAP 19  
 DB |||||  
 1 TPEKEPTAAPAPAPAP 19

RESULT 5  
 AAU01433  
 ID AAU01433 standard; peptide; 18 AA.  
 XX  
 AC AAU01433;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide #19 derived from natural latex rubber allergen Hev b 5.  
 XX Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 KW Hevea brasiliensis.  
 XX WO200123429-A1.  
 XX  
 PN 05-APR-2001.  
 XX  
 PD 22-SEP-2000; 2000WO-AU001182.  
 XX  
 PF 24-SEP-1999; 99AU-00003057.  
 PR 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 DR New peptide and immunospecific antibodies, useful for treatment and/or  
 XX prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 PT Disclosure; Fig 10; 101pp; English.  
 PS The sequence represents the amino acid sequence of peptide #19 derived  
 XX from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 18 AA;  
 SQ

PS Disclosure; Fig 10; 101pp; English.  
 XX The sequence represents the amino acid sequence of peptide #19 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 18 AA;  
 SQ

Query Match 88.8%; Score 95; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. NO. 0.00019;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAPAPAP 18  
 DB |||||  
 1 TPEKEPTAAPAPAPAP 18

RESULT 6  
 AAU01445  
 ID AAU01445 standard; peptide; 18 AA.  
 XX  
 AC AAU01445;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide #31 derived from natural latex rubber allergen Hev b 5.  
 XX Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 KW Hevea brasiliensis.  
 XX WO200123429-A1.  
 XX  
 PN 05-APR-2001.  
 XX  
 PD 22-SEP-2000; 2000WO-AU001182.  
 XX  
 PF 24-SEP-1999; 99AU-00003057.  
 PR 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 DR New peptide and immunospecific antibodies, useful for treatment and/or  
 XX prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 PT Disclosure; Fig 10; 101pp; English.  
 PS The sequence represents the amino acid sequence of peptide #31 derived  
 XX from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 18 AA;  
 SQ

Query Match 88.8%; Score 95; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. NO. 0.00019;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.00019; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKEPTAAPAEPEAPE 20  
 |||||  
 Db 1 EKEPTAAPAEPEAPE 18

RESULT 7  
 AAU01434  
 ID AAU01434 standard; peptide; 17 AA.  
 AC AAU01434;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide #20 derived from natural latex rubber allergen Hev b 5.  
 XX  
 KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX  
 OS Hevea brasiliensis.  
 XX  
 PN WO200123429-A1.  
 XX  
 PD PD  
 XX  
 PF AAU01434;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide #20 derived from natural latex rubber allergen Hev b 5.  
 XX  
 KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX  
 OS Hevea brasiliensis.  
 XX  
 PN WO200123429-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 22-SEP-2000; 2000WO-AU001182.  
 XX  
 PR 24-SEP-1999; 99AU-00003057.  
 PR 24-JUL-2000; 2000AU-00008964.  
 XX  
 PA (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX  
 PI O'hehir R;  
 XX  
 DR WPI; 2001-258129/26.  
 XX  
 PT New peptide and immunospecific antibodies, useful for treatment and/or  
 PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 derived from or homologous to Hev b 5.  
 XX  
 PS Disclosure; Fig 10; 101pp; English.  
 XX  
 CC The sequence represents the amino acid sequence of peptide #20 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX  
 SQ Sequence 17 AA;

Query Match 85.0%; Score 91; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.00053;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAP 17  
 |||||  
 Db 1 TPEKEPTAAPAEPEAP 17

RESULT 8  
 AAU01446  
 ID AAU01446 standard; peptide; 17 AA.  
 AC AAU01446;  
 XX  
 DT 18-JUL-2001 (first entry)

XX Peptide #32 derived from natural latex rubber allergen Hev b 5.  
 DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 XX latex hypersensitivity.  
 KW Hevea brasiliensis.  
 OS  
 XX WO200123429-A1.  
 PN  
 XX 05-APR-2001.  
 PD  
 XX  
 PF 22-SEP-2000; 2000WO-AU001182.  
 XX  
 PR 24-SEP-1999; 99AU-00003057.  
 PR 24-JUL-2000; 2000AU-00008964.  
 XX  
 PA (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX  
 PI O'hehir R;  
 XX  
 DR WPI; 2001-258129/26.  
 XX  
 PT New peptide and immunospecific antibodies, useful for treatment and/or  
 PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 derived from or homologous to Hev b 5.  
 XX  
 PS Disclosure; Fig 10; 101pp; English.  
 XX  
 CC The sequence represents the amino acid sequence of peptide #32 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX  
 SQ Sequence 17 AA;

Query Match 84.1%; Score 90; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.00069;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEEPTAAPAEPEAPE 20  
 |||||  
 Db 1 KEEPTAAPAEPEAPE 17

RESULT 9  
 AAU01447  
 ID AAU01447 standard; peptide; 16 AA.  
 XX  
 AC AAU01447;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide #33 derived from natural latex rubber allergen Hev b 5.  
 XX  
 KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX  
 OS Hevea brasiliensis.  
 XX  
 PN WO200123429-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 22-SEP-2000; 2000WO-AU001182.  
 XX  
 PR 24-SEP-1999; 99AU-00003057.



QY 6 EPTAAPAEPEAPE 20  
 DB 1 EPTAAPAEPEAPE 15  
 RESULT 12  
 ID AAU01436 standard; peptide; 15 AA.  
 XX AC AAU01436;  
 XX DT 18-JUL-2001 (first entry)  
 XX Peptide #22 derived from natural latex rubber allergen Hev b 5.  
 DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX Hevea brasiliensis.  
 OS WO200123429-A1.  
 PN 22-SEP-2000; 2000WO-AU001182.  
 XX 24-SEP-1999; 99AU-00003057.  
 XX 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 XX New peptide and immunospecific antibodies, useful for treatment and/or  
 PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 XX Disclosure; Fig 10; 101pp; English.  
 XX The sequence represents the amino acid sequence of peptide #22 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 15 AA;  
 Query Match 74.8%; Score 80; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0088;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPEKEEPTAAPAPE 15  
 DB 1 TPEKEEPTAAPAPE 15  
 RESULT 13  
 AAU01437  
 ID AAU01437 standard; peptide; 14 AA.  
 XX AC AAU01437;  
 XX DT 18-JUL-2001 (first entry)  
 XX Peptide #23 derived from natural latex rubber allergen Hev b 5.  
 DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX Hevea brasiliensis.  
 OS WO200123429-A1.  
 PN 22-SEP-2000; 2000WO-AU001182.  
 XX 24-SEP-1999; 99AU-00003057.  
 XX 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 XX New peptide and immunospecific antibodies, useful for treatment and/or  
 PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 XX Disclosure; Fig 10; 101pp; English.  
 XX The sequence represents the amino acid sequence of peptide #22 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 14 AA;  
 Query Match 70.1%; Score 75; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPEKEEPTAAPAPE 14  
 DB 1 TPEKEEPTAAPAPE 14  
 RESULT 14  
 AAU01449  
 ID AAU01449 standard; peptide; 14 AA.  
 XX AC AAU01449;  
 XX DT 18-JUL-2001 (first entry)  
 XX Peptide #35 derived from natural latex rubber allergen Hev b 5.  
 DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX Hevea brasiliensis.  
 OS WO200123429-A1.  
 PN 22-SEP-2000; 2000WO-AU001182.  
 XX 24-SEP-1999; 99AU-00003057.  
 XX 24-JUL-2000; 2000AU-00008964.

XX Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX Hevea brasiliensis.  
 OS WO200123429-A1.  
 PN 22-SEP-2000; 2000WO-AU001182.  
 XX 24-SEP-1999; 99AU-00003057.  
 XX 24-JUL-2000; 2000AU-00008964.  
 XX (MONU ) UNIV MONASH.  
 PA (REAS-) COOP RES CENT ASTHMA.  
 XX O'hehir R;  
 XX WPI; 2001-258129/26.  
 XX New peptide and immunospecific antibodies, useful for treatment and/or  
 PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
 PT derived from or homologous to Hev b 5.  
 XX Disclosure; Fig 10; 101pp; English.  
 XX The sequence represents the amino acid sequence of peptide #23 derived  
 CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
 CC The peptide is capable of interacting with T cells and modifying T cell  
 CC function when incubated with cells from subjects with a condition  
 CC characterised by an aberrant, unwanted or inappropriate immune response  
 CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
 CC the manufacture of a medicament for the treatment and/or prophylaxis of  
 CC latex hypersensitivity in a subject. They are also useful for diagnosing  
 CC or monitoring latex hypersensitivity in a subject  
 XX Sequence 14 AA;  
 Query Match 70.1%; Score 75; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPEKEEPTAAPAPE 14  
 DB 1 TPEKEEPTAAPAPE 14  
 RESULT 14  
 AAU01449  
 ID AAU01449 standard; peptide; 14 AA.  
 XX AC AAU01449;  
 XX DT 18-JUL-2001 (first entry)  
 XX Peptide #35 derived from natural latex rubber allergen Hev b 5.  
 DE Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
 KW latex hypersensitivity.  
 XX Hevea brasiliensis.  
 OS WO200123429-A1.  
 PN 22-SEP-2000; 2000WO-AU001182.  
 XX 24-SEP-1999; 99AU-00003057.  
 XX 24-JUL-2000; 2000AU-00008964.

|           |  |
|-----------|--|
| PA        | (MONU ) UNIV MONASH.   |
| PA        | (REAS-) COOP RES CENT ASTHMA.  |
| XX        | O'hehir R;   |
| XX        | WPI; 2001-258129/26.   |
| DR        |  |
| XX        | New peptide and immunospecific antibodies, useful for treatment and/or   |
| PT        | prophylaxis of latex hypersensitivity, comprises an amino acid sequence  |
| PT        | derived from or homologous to Hev b 5.                                   |
| PT        |  |
| XX        | Disclosure; Fig 10; 101pp; English.                                      |
| PS        |  |
| XX        | The sequence represents the amino acid sequence of peptide #35 derived   |
| CC        | from natural latex rubber allergen Hev b 5, isolated from rubber plant.  |
| CC        | The peptide is capable of interacting with T cells and modifying T cell  |
| CC        | function when incubated with cells from subjects with a condition        |
| CC        | characterised by an aberrant, unwanted or inappropriate immune response  |
| CC        | to Hev b 5. The peptides, or antibodies to these peptides are useful for |
| CC        | the manufacture of a medicament for the treatment and/or prophylaxis of  |
| CC        | latex hypersensitivity in a subject. They are also useful for diagnosing |
| CC        | or monitoring latex hypersensitivity in a subject                        |
| XX        |  |
| SQ        | Sequence 14 AA;  |
|           | Query Match 70.1%; Score 75; DB 4; Length 14;                            |
|           | Best Local Similarity 100.0%; Pred.No. 0.031;                            |
|           | Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;              |
| Qy        | 7 PTAAPAEPEAPE 20<br>                   <br>1 PTAAPAEPEAPE 14            |
| Dd        |  |
| RESULT 15 |  |
| AAU01438  |  |
| ID        | AAU01438 standard; peptide; 13 AA.                                       |
| XX        |  |
| AC        | AAU01438;  |
| XX        |  |
| DT        | 18-JUL-2001 (first entry)  |
| XX        |  |
| DE        | Peptide #24 derived from natural latex rubber allergen Hev b 5.          |
| XX        |  |
| KW        | Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  |
| KW        | latex hypersensitivity.  |
| XX        |  |
| OS        | Hevea brasiliensis.  |
| XX        |  |
| PN        | WO200123429-A1.  |
| XX        |  |
| PD        | 05-APR-2001.   |
| XX        |  |
| PF        | 22-SEP-2000; 2000MO-AU001182.  |
| XX        |  |
| PR        | 24-SEP-1999; 99AU-00003057.  |
| PR        | 24-JUL-2000; 2000AU-00008964.  |
| XX        |  |
| FA        | (MONU ) UNIV MONASH.   |
| PA        | (REAS-) COOP RES CENT ASTHMA.  |
| XX        |  |
| PI        | O'hehir R;   |
| XX        |  |
| DR        | WPI; 2001-258129/26.   |
| XX        |  |
| PT        | New peptide and immunospecific antibodies, useful for treatment and/or   |
| PT        | prophylaxis of latex hypersensitivity, comprises an amino acid sequence  |
| PT        | derived from or homologous to Hev b 5.                                   |
| XX        |  |
| FS        | Disclosure; Fig 10; 101pp; English.                                      |
| XX        |  |
| CC        | The sequence represents the amino acid sequence of peptide #24 derived   |
| CC        | from natural latex rubber allergen Hev b 5, isolated from rubber plant.  |

```

Db      1 TPAPEPEAPE 13
|||||
RESULT 17
AAU01439
ID AAU01439 standard; peptide; 12 AA.
XX AC AAU01439;
XX XX
DT 18-JUL-2001 (first entry)
XX XX
DE Peptide #25 derived from natural latex rubber allergen Hev b 5.
XX Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;
KW latex hypersensitivity.
XX Hevea brasiliensis.
OS
PN WO200123429-A1.
XX
PF 22-SEP-2000; 2000WO-AU001182.
XX
PR 24-SEP-1999; 99AU-00003057.
PR 24-JUL-2000; 2000AU-00008964.
XX
PA (MONU ) UNIV MONASH.
PA (REAS-) COOP RES CENT ASTHMA.
XX
PI O'hehir R;
XX
DR WPI; 2001-258129/26.
XX
PT New peptide and immunospecific antibodies, useful for treatment and/or
PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence
PT derived from or homologous to Hev b 5.
XX
PS Disclosure; Fig 10; 101pp; English.
XX
CC The sequence represents the amino acid sequence of peptide #25 derived
CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.
CC The peptide is capable of interacting with T cells and modifying T cell
CC function when incubated with cells from subjects with a condition
CC characterised by an aberrant, unwanted or inappropriate immune response
CC to Hev b 5. The peptides, or antibodies to these peptides are useful for
CC the manufacture of a medicament for the treatment and/or prophylaxis of
CC latex hypersensitivity in a subject. They are also useful for diagnosing
CC or monitoring latex hypersensitivity in a subject
XX
SQ Sequence 12 AA;
Query Match 58.9%; Score 63; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPEKEEPTAAPA 12
Db 1 TPEKEEPTAAPA 12
|||||
RESULT 18
AAU01451
ID AAU01451 standard; peptide; 12 AA.
XX AC AAU01451;
XX XX
DT 18-JUL-2001 (first entry)
XX XX
DE Peptide #37 derived from natural latex rubber allergen Hev b 5.
XX Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;
KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;

```

```

KW latex hypersensitivity.
XX
OS Hevea brasiliensis.
XX
PN WO200123429-A1.
XX
PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000WO-AU001182.
XX
PR 24-SEP-1999; 99AU-00003057.
PR 24-JUL-2000; 2000AU-00008964.
XX
PA (MONU ) UNIV MONASH.
PA (REAS-) COOP RES CENT ASTHMA.
XX
PI O'hehir R;
XX
DR WPI; 2001-258129/26.
XX
PT New peptide and immunospecific antibodies, useful for treatment and/or
PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence
PT derived from or homologous to Hev b 5.
XX
PS Disclosure; Fig 10; 101pp; English.
XX
CC The sequence represents the amino acid sequence of peptide #37 derived
CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.
CC The peptide is capable of interacting with T cells and modifying T cell
CC function when incubated with cells from subjects with a condition
CC characterised by an aberrant, unwanted or inappropriate immune response
CC to Hev b 5. The peptides, or antibodies to these peptides are useful for
CC the manufacture of a medicament for the treatment and/or prophylaxis of
CC latex hypersensitivity in a subject. They are also useful for diagnosing
CC or monitoring latex hypersensitivity in a subject
XX
SQ Sequence 12 AA;
Query Match 58.9%; Score 63; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 AAPAEPEAPE 20
Db 1 AAPAEPEAPE 12
|||||
RESULT 19
AAV53666
ID AAV53666 standard; protein; 4412 AA.
XX AC AAV53666;
XX
DT 22-FEB-2000 (first entry)
XX
DE Sequence gi/1017427/emb/CAA62189 from an alignment with protein 608.
XX
KW Mechanical stress; gene therapy; protein 608; osteoporosis; bone density;
KW bone development; gi/1017427/emb/CAA62189.
XX
OS Unidentified.
XX
PN WO9960164-A1.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-US011066.
XX
PR 15-MAY-1998; 98US-0085673P.
XX
PA (QUAR-) QUARK BIOTECH INC.
XX
PI Binat P, Mor O, Skalter R, Feinstein E, Faerman A;

```

XX WPI; 2000-053304/04.  
 XX Identification of stress induced genes for determining risk and  
 XX preventing, treating or controlling osteoporosis.  
 PT Claim 32; Fig 6A-R; 308pp; English.  
 XX  
 XX The present sequence is obtained from a clustral X alignment with protein  
 CC 608. Protein 608 was identified using the method of the invention after  
 CC subjecting rat osteoblasts to mechanical stress. Expression of the 608  
 CC gene was found to be upregulated by about 3-fold in cells subjected to  
 CC mechanical strain. The specification describes a method for the  
 CC identification of genes responsive to a specific mechanical stress. The  
 CC method comprises applying the mechanical stress to an organism (tissue or  
 CC cells comprising bone cells), isolating the specific cellular fractions  
 CC and extracting mRNA from them, and differentially analysing the mRNA in  
 CC comparison with control samples. The method is used to identify genes  
 CC whose expression is responsive to a specific stress. The identified genes  
 CC are employed in determining risk associated with a physiological or  
 CC disease state. The risk determination methods are used for testing a  
 CC medicament for gene therapy. These medicaments, or genes identified by  
 CC the method of the invention, are used for treating, preventing or  
 CC controlling a physiological or disease state (especially osteoporosis or  
 CC bone density or other factors causing or contributing to osteoporosis or  
 CC its symptoms or other conditions involved in mechanical stress or its  
 CC lack. The methods can also be used for advancing research or studies in  
 CC bone development  
 XX  
 XX Sequence 4412 AA;  
 SQ  
 Query Match 58.9%; Score 63; DB 3; Length 4412;  
 Best Local Similarity 57.9%; Pred. No. 1.9e+02;  
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 PEKEEPTAAPAPPEAPE 20  
 |||: |||: |||: |||:  
 Db 3677 PEKKVPEAIPKPEPPPE 3695  
 RESULT 20  
 ADQ89964  
 ID ADQ89964 standard; protein; 34350 AA.  
 XX  
 AC ADQ89964;  
 XX  
 XX 21-OCT-2004 (first entry)  
 DT  
 XX Antagonist of cell cycle progression polypeptide #197.  
 DE  
 XX Cytostatic; cancer; cell division cycle; mitosis; meiosis;  
 KW cell cycle progression.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2004063362-A2.  
 PN  
 XX 29-JUL-2004.  
 PD  
 XX 31-DEC-2003; 2003WO-GB005635.  
 XX  
 XX 10-JAN-2003; 2003US-0439123P.  
 PR  
 XX 06-MAY-2003; 2003US-0468402P.  
 PR  
 XX (CYCL-) CYCLACEL LTD.  
 XX  
 XX Glover D, Bell G, Frenz L, Midgley C;  
 PI  
 XX WPI; 2004-544089/52.  
 DR  
 XX N-PSDB; ADQ89963.  
 DR  
 XX New cell cycle progression genes and proteins for modulating cell cycle  
 PT progression in cells, for preventing, treating or diagnosing cell

PT proliferative diseases (e.g. cancer) or for identifying modulators of  
 PT mitosis or meiosis.  
 XX  
 XX Claim 2; SEQ ID NO 394; 461pp; English.  
 PS  
 XX The present invention relates to a polynucleotide for preventing,  
 CC treating or diagnosing a disease in an individual. The composition or the  
 CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for  
 CC diagnosing, preventing or treating diseases (e.g. cell proliferative  
 CC diseases such as cancer) in an individual. These may also be used for  
 CC identifying substances capable of binding to or modulating the function  
 CC of the polypeptide, capable of affecting the function of the  
 CC corresponding gene, or capable of inhibiting the cell division cycle or  
 CC cell cycle progression, preferably mitosis and/or meiosis. The present  
 CC sequence represents an antagonist of cell cycle progression protein  
 CC sequence.  
 XX  
 XX Sequence 34350 AA;  
 SQ  
 Query Match 58.9%; Score 63; DB 8; Length 34350;  
 Best Local Similarity 57.9%; Pred. No. 1.3e+03;  
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 PEKEEPTAAPAPPEAPE 20  
 |||: |||: |||: |||:  
 Db 11489 PEKKVPEAIPKPEPPPE 11507  
 RESULT 21  
 AAW72911  
 ID AAW72911 standard; protein; 666 AA.  
 XX  
 AC AAW72911;  
 XX  
 XX 21-JAN-1999 (first entry)  
 DT  
 XX Mycobacterium tuberculosis antigen RD1-ORF9B.  
 DE  
 XX Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;  
 KW infection.  
 KW  
 XX Mycobacterium tuberculosis.  
 OS  
 XX WO9844119-A1.  
 PN  
 XX 08-OCT-1998.  
 PD  
 XX 01-APR-1998; 98WO-DK000132.  
 PF  
 XX 02-APR-1997; 97DK-00000376.  
 PR  
 XX 18-APR-1997; 97US-0044624P.  
 PR  
 XX 10-NOV-1997; 97DK-00001277.  
 PR  
 XX 03-JAN-1998; 98US-0070488P.  
 PR  
 XX (STAT-) STATENS SERUM INST.  
 PA  
 XX Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB;  
 PI Oettinger T, Florio W;  
 PI  
 XX WPI; 1998-542705/46.  
 DR  
 XX N-PSDB; AAV63935.  
 DR  
 XX New isolated mycobacteria polypeptides and nucleic acids - used for  
 PT developing products for the diagnosis of or vaccination against  
 PT mycobacterial infections, particularly tuberculosis.  
 PT  
 XX Claim 1; Page 183-185; 163pp; English.  
 PS  
 XX The present sequence represents a Mycobacterium tuberculosis protein.  
 CC Products from the present invention, which describes protein fragments  
 CC and nucleic acid fragments derived from M.tuberculosis, can be used in  
 CC the detection of and prevention of mycobacterial infections. In  
 CC particular, the proteins and nucleic acids can be used for the diagnosis

CC of or vaccination against tuberculosis caused by *M. tuberculosis*, *M.*  
CC africanum or *M. bovis*  
XX  
SQ Sequence 666 AA;

Query Match 57.0%; Score 61; DB 2; Length 666;  
Best Local Similarity 55.6%; Pred. No. 53;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAARPEAPAP 19  
| : ||| : ||| : |||  
Db 156 PTPQTGTGAPOQPESAP 173

RESULT 22  
AAY21928  
ID AAY21928 standard; protein; 666 AA.  
AC AAY21928;  
XX  
DT 06-SEP-1999 (first entry)  
XX  
DE Amino acid sequence of antigen RD1-ORF9B.  
XX  
KW Immunogenic; Mycobacterium tuberculosis; immune response; infection;  
KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;  
KW pharmaceutical; vaccination; *M. africanum*; *M. bovis*; CFP7A; CFP30A;  
KW CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;  
KW CFP25A; CFP30B; CFP7B.  
XX  
OS Mycobacterium tuberculosis.  
XX  
XX  
PN WO924577-A1.  
XX  
PD 20-MAY-1999.  
XX  
PF 08-OCT-1998; 98WO-DK000438.  
XX  
PR 10-NOV-1997; 97DK-00001277.  
PR 05-JAN-1998; 98US-0070488P.  
PR 01-APR-1998; 98WO-DK000132.  
XX  
PA (STAT-) STATENS SERUM INST.  
XX  
PI Andersen P, Skjot R;  
XX  
DR WPI; 1999-347282/29.  
DR N-PSDB; AAX81042.  
XX  
PT New immunogenic fragment of *Mycobacterium tuberculosis*.  
XX  
PS Example 2; Page 202-204; 265pp; English.  
XX  
CC The invention describes a substantially pure immunogenic polypeptide  
CC fragment (I) from *Mycobacterium tuberculosis* that is able to evoke a  
CC protective immune response against infections by mycobacteria belonging  
CC to the tuberculosis complex. The invention provides a (1) fusion  
CC polypeptide comprising at least one polypeptide fragment (I) and at least  
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell  
CC epitope from *M. tuberculosis* protein ESAT-6, or MPT59 and a second  
CC different amino acid sequence from *M. tuberculosis*, and/or including a  
CC sequence which protects the first amino acid sequence from *in vivo*  
CC degradation or post-translational processing; (3) a nucleic acid fragment  
CC that encodes the above polypeptides. The polypeptides and nucleic acid  
CC are useful as pharmaceuticals, for diagnosis of and as antigens for  
CC vaccination against TB caused by *Mycobacterium tuberculosis*, *africanum* or  
CC *bovis*. The polypeptides are also useful for diagnosing ongoing or  
CC previous sensitization in an animal with bacteria belonging to the  
CC tuberculosis complex. The invention also describes the use of CFP7A or  
CC CFP30A or a T-cell epitope of for the induction of a strong immune  
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell  
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin  
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,

CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-  
CC cell epitope of for the preparation of an immunological composition; and  
CC for the preparation of a subunit vaccine  
XX  
SQ Sequence 666 AA;

Query Match 57.0%; Score 61; DB 2; Length 666;  
Best Local Similarity 55.6%; Pred. No. 53;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAARPEAPAP 19  
| : ||| : ||| : |||  
Db 156 PTPQTGTGAPOQPESAP 173

RESULT 23  
AAB19846  
ID AAB19846 standard; protein; 666 AA.  
XX  
AC AAB19846;  
XX  
DT 05-MAR-2001 (first entry)  
XX  
DE Mycobacterium tuberculosis protein MTBNS.  
XX  
KW MTBNS; tuberculosis; BCG; vaccine; infection; diagnosis.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO200066157-A1.  
XX  
PD 09-NOV-2000.  
XX  
PF 04-MAY-2000; 2000WO-US012257.  
XX  
PR 04-MAY-1999; 99US-0132505P.  
XX  
PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.  
XX  
PI Gennaro ML;  
XX  
DR WPI; 2001-007153/01.  
DR N-PSDB; AAA89039.  
XX  
XX Novel polypeptide encoded by open reading frames present in *Mycobacterium*  
PT tuberculosis genome and not by the BCG strain of *M. bovis*, useful as  
PT vaccine and for diagnosing tuberculosis infection.  
XX  
PS Claim 11; Fig 1; 35pp; English.  
XX  
CC The present sequence is that of the *Mycobacterium tuberculosis* MTBNS  
CC protein. This is 1 of 8 proteins, i.e. MTBNS-8 (see AAB19842-49), encoded  
CC by 8 open reading frames (see AAA89035-42) identified as being present in  
CC the genome of *M. tuberculosis* but absent from the genome of the BCG  
CC strain of *Mycobacterium bovis*. MTBNS-8 represent reagents that are useful  
CC in discriminating between *M. tuberculosis* and BCG and, in particular, for  
CC diagnostic methods which discriminate between exposure of a subject to *M.*  
CC tuberculosis and vaccination with BCG. The invention features these MTBNS  
CC polypeptides, functional fragments of them, DNA encoding them, vectors,  
CC transformed cells, and diagnostic, therapeutic, and prophylactic  
CC (vaccine) methods, including genetic vaccination methods  
XX  
SQ Sequence 666 AA;

Query Match 57.0%; Score 61; DB 4; Length 666;  
Best Local Similarity 55.6%; Pred. No. 53;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PEKEPTAARPEAPAP 19  
| : ||| : ||| : |||  
Db 156 PTPQTGTGAPOQPESAP 173



```
RESULT 24
ABU05989
ID ABU05989 standard; protein; 666 AA.
XX
AC ABU05989;
XX
DT 08-APR-2003 (first entry)
XX
DE M. tuberculosis and M. leprae marker protein #640.
XX
KW Mycobacterioses; survival; virulence; protective antigen; vaccine;
KW mycobacterial disease; tuberculosis; leprosy.
XX
OS Mycobacterium tuberculosis.
OS Mycobacterium leprae.
XX
PN WO200274903-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-IB001973.
XX
PR 22-FEB-2001; 2001US-0270123P.
XX
PA (INSP ) INST PASTEUR.
XX
PI Cole S;
XX
PS WPI; 2002-759885/82.
XX
PT Identifying and selecting genes for survival or virulence of mycobacteria
PT by a comparative genomic analysis of the sequences of Mycobacterium
PT tuberculosis and M. leprae.
XX
PS Claim 17; Page 864-865; 874pp; English.
XX
CC This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a marker protein from
CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
CC method of the invention
XX
SQ Sequence 666 AA;
Query Match 57.0%; Score 61; DB 5; Length 666;
Best Local Similarity 55.6%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 PEKEEPTAAPAEPEAPAP 19
Db 156 PTPQTPTGAPQOPEAPAP 173
RESULT 25
ABU37020
ID ABU37020 standard; protein; 666 AA.
XX
AC ABU37020;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #22547.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
```

```
XX
OS Mycobacterium tuberculosis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR N-PSDB; ACA40890.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 64944; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 666 AA;
Query Match 57.0%; Score 61; DB 6; Length 666;
Best Local Similarity 55.6%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 PEKEEPTAAPAEPEAPAP 19
Db 156 PTPQTPTGAPQOPEAPAP 173
RESULT 26
```





CC latex hypersensitivity in a subject. They are also useful for diagnosing  
CC or monitoring latex hypersensitivity in a subject  
XX  
SQ Sequence 20 AA;

Query Match 55.1%; Score 59; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAP 11  
|||  
Db 10 TPEKEPTAAP 20

## RESULT 31

AAU01422  
ID AAU01422 standard; peptide; 20 AA.

AC AAU01422;

DT 18-JUL-2001 (first entry)

DE Peptide #8 derived from natural latex rubber allergen Hev b 5.

KW Rubber plant; Hev b 5; allergen; natural rubber latex; allergy; T cell;  
XX latex hypersensitivity.

OS Hevea brasiliensis.

PN WO200123429-A1.

PD 05-APR-2001.

PF 22-SEP-2000; 2000WO-AU001182.

PR 24-SEP-1999; 99AU-00003057.

PR 24-JUL-2000; 2000AU-00008964.

PA (MONU ) UNIV MONASH.

PA (REAS-) COOP RES CENT ASTHMA.

O'hehir R;

WPI; 2001-258129/26.

PT New peptide and immunospecific antibodies, useful for treatment and/or  
PT prophylaxis of latex hypersensitivity, comprises an amino acid sequence  
PT derived from or homologous to Hev b 5.

PS Claim 8; Fig 1; 101pp; English.

CC The sequence represents the amino acid sequence of peptide #8 derived  
CC from natural latex rubber allergen Hev b 5, isolated from rubber plant.  
CC The peptide is capable of interacting with T cells and modifying T cell  
CC function when incubated with cells from subjects with a condition  
CC characterised by an aberrant, unwanted or inappropriate immune response  
CC to Hev b 5. The peptides, or antibodies to these peptides are useful for  
CC the manufacture of a medicament for the treatment and/or prophylaxis of  
CC latex hypersensitivity in a subject. They are also useful for diagnosing  
CC or monitoring latex hypersensitivity in a subject

SQ Sequence 20 AA;

Query Match 55.1%; Score 59; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 APAEPEAPE 20  
|||  
Db 1 APAEPEAPE 11

## RESULT 32

ABB84377

ID ABB84377 standard; peptide; 20 AA.

AC ABB84377;

DT 04-NOV-2002 (first entry)

DE H. brasiliensis associated peptide SEQ ID 24.

KW BPI; bifunctional peptide inhibitor; epitope; TCR; T cell receptor;  
KW Signal-2; immunomodulatory; type-2 immunity.

OS Hevea brasiliensis.

PN WO200250250-A2.

PD 27-JUN-2002.

PF 17-DEC-2001; 2001WO-US048632.

PR 18-DEC-2000; 2000US-00739466.

PA (UNIV ) UNIV KANSAS.

PI Murray JS, Siahaan TJ, Hu Y;

XX WPI; 2002-557657/59.

CC Bifunctional peptide inhibitors comprising a first portion derived from a  
CC peptide recognition receptor epitope and a second portion derived from a  
CC Signal-2 moiety, useful for modulating immune response.

PS Claim 8; Page 97; 108pp; English.

CC This invention describes a novel peptide comprising a first portion  
CC having a sequence at least 10% homologous to a peptide derived from a  
CC peptide recognition receptor (TCR) epitope capable of initiating a signal  
CC in a T-cell, and a second portion having a sequence at least 10%  
CC homologous to a peptide derived from a Signal-2 moiety capable of  
CC initiating a second signal in a T-cell. The resulting peptide has  
CC immunomodulatory activity and are used for constructing bifunctional  
CC peptide inhibitors (BPI) by selecting a known health condition,  
CC identifying a TCR epitope specific for the health condition, and  
CC incorporating the peptide recognition receptor (TCR) epitope into the  
CC BPI. The peptide is useful for replacing type-1 immunity with type-2  
CC immunity, or type-2 immunity to type-1 immunity, and for modulating  
CC immune response. ABB84354-ABB84399 represent BPI associated peptides  
CC useful to the invention

SQ Sequence 20 AA;

Query Match 55.1%; Score 59; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPEKEPTAAP 11  
|||  
Db 10 TPEKEPTAAP 20

## RESULT 33

ABG97431

ID ABG97431 standard; protein; 382 AA.

AC ABG97431;

DT 16-DEC-2002 (first entry)

DE S. ghanaensis UNBU protein.

KW Enediye; warhead structure; cancer; cytostatic; PKSE; enzyme;  
KW polyketide synthase; TBBC; UNBU; UNBU; genetic engineering;  
KW thioesterase unique to enediye biosynthetic loci;



XX DE Human DITHP polypeptide #177.  
 XX KW Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;  
 KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;  
 KW inflammatory disorder; viral infection; bacterial infection; seizure;  
 KW fungal infection; parasitic infections; developmental disorder; breast;  
 KW endocrine disorder; metabolic disorder; neurological disorder; cervix;  
 KW gastrointestinal disorder; transport disorder; gene therapy; kidney;  
 KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;  
 KW thymus.  
 XX OS Homo sapiens.  
 XX FN WO200220754-A2.  
 XX PD 14-MAR-2002.  
 XX PF 29-AUG-2001; 2001WO-US027127.  
 XX PR 05-SEP-2000; 2000US-0229747P.  
 PR 05-SEP-2000; 2000US-0229748P.  
 PR 05-SEP-2000; 2000US-0229749P.  
 PR 05-SEP-2000; 2000US-0229750P.  
 PR 05-SEP-2000; 2000US-0229751P.  
 PR 05-SEP-2000; 2000US-0230583P.  
 PR 06-SEP-2000; 2000US-0230505P.  
 PR 06-SEP-2000; 2000US-0230514P.  
 PR 06-SEP-2000; 2000US-0230515P.  
 PR 06-SEP-2000; 2000US-0230517P.  
 PR 06-SEP-2000; 2000US-0230518P.  
 PR 06-SEP-2000; 2000US-0230519P.  
 PR 06-SEP-2000; 2000US-0230595P.  
 PR 06-SEP-2000; 2000US-0230597P.  
 PR 06-SEP-2000; 2000US-0230598P.  
 PR 06-SEP-2000; 2000US-0230599P.  
 PR 06-SEP-2000; 2000US-0230610P.  
 PR 06-SEP-2000; 2000US-0230865P.  
 PR 06-SEP-2000; 2000US-0230988P.  
 PR 07-SEP-2000; 2000US-0230951P.  
 PR 07-SEP-2000; 2000US-0231163P.  
 PR 07-SEP-2000; 2000US-0231167P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;  
 PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;  
 PI Momiya MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;  
 PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;  
 PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;  
 XX WI; 2002-383054/41.  
 DR N-PSDB; ABK71710.  
 XX An isolated polynucleotide useful in diagnostics and therapeutics.  
 XX PS Claim 29; Page 632-633; 686pp; English.  
 XX The invention relates to human diagnostic and therapeutic (dithp)  
 CC polynucleotides and their associated polypeptides (DITHP polypeptides).  
 CC The sequences of the invention are used in the treatment and diagnosis of  
 CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers  
 CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,  
 CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or  
 CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,  
 CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal  
 CC infections, parasitic infections, developmental disorders (e.g. anaemia,  
 CC epilepsies), seizure disorders (e.g. cerebral palsy, spina bifida),  
 CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders  
 CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,  
 CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal  
 CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders  
 CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences

CC ABG59943-ABG60220 represent human DITHP polypeptides of the invention  
 XX SQ Sequence 267 AA;  
 Query Match 53.3%; Score 57; DB 5; Length 267;  
 Best Local Similarity 62.5%; Pred. No. 64;  
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 PEKKEPTAAPAPPEAP 17  
 |||: |||||  
 Db 222 PEKKAPVAPPEPEVP 237  
 RESULT 36  
 ABP25792  
 ID ABP25792 standard; protein; 379 AA.  
 XX AC ABP25792;  
 XX DT 02-JUL-2002 (first entry)  
 XX Streptococcus polypeptide SEQ ID NO 760.  
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus pyogenes.  
 XX WO200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB004789.  
 XX PR 27-OCT-2000; 2000GB-00026333.  
 PR 24-NOV-2000; 2000GB-00028727.  
 PR 07-MAR-2001; 2001GB-00005640.  
 XX (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX WI; 2002-352536/38.  
 DR N-PSDB; ABN66423.  
 XX New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.  
 XX Claim 1; Page 3227; 4525pp; English.  
 XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (Si), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX Sequence 379 AA;

CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 379 AA;

Query Match 53.3%; Score 57; DB 6; Length 379;  
Best Local Similarity 50.0%; Pred. No. 89;  
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPEKEPTAAPAEPEAPE 20  
:|||||:|:|:|:  
Db 221 SPESEEPSVAAPSEETPSE 240

RESULT 38  
ADR83966  
ID ADR83966 standard; protein; 379 AA.  
XX  
AC ADR83966;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE *S. pyogenes* hyperimmune system reactive antigen Spy2009.  
XX  
KW hyperimmune serum reactive antigen; vaccine; anticaline.  
XX  
OS Streptococcus pyogenes.  
XX  
PN WO2004078907-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-EP002087.  
XX  
PR 04-MAR-2003; 2003EP-00450061.  
XX  
PA (INTE-) INTERCELL AG.  
XX  
PI Meinke A, Nagy E, Winkler B, Gelbmann D;  
XX  
XX WPI; 2004-653698/63.  
DR N-PSDB; ADR83816.  
XX  
XX New isolated nucleic acid molecules encoding hyperimmune serum-reactive  
PT antigens from Streptococcus pyogenes, useful for diagnosing, preventing  
PT and treating *S. pyogenes* infections.  
PS Claim 13; SEQ ID NO 234; 145pp; English.  
XX  
XX This invention describes a novel nucleic acid molecule encoding a  
CC hyperimmune serum reactive antigen or its fragment from Streptococcus  
CC pyogenes. The nucleic acid molecule or hyperimmune serum-reactive antigen  
CC or its fragment are useful for the manufacture of a pharmaceutical  
CC preparation, especially a vaccine, against *S. pyogenes* infection. In  
CC addition, the hyperimmune serum reactive antigen or fragment is used for  
CC the isolation and/or purification and/or identification of an interaction  
CC partner of the hyperimmune serum reactive antigen or its fragment, for  
CC the generation of a peptide (e.g. anticalines) binding to the antigen or  
CC fragment, or for the manufacture of a functional nucleic acid selected  
CC from aptamers and spiegelmers. The nucleic acid molecule may also be used  
CC for the manufacture of functional ribonucleic acids, such as ribozymes,  
CC antisense nucleic acids and siRNA. ADR83733-ADR84189 represent *S.*  
CC *pyogenes* hyperimmune serum reactive antigens, fragments and the encoding  
CC polynucleotide described in the invention.  
XX  
XX Sequence 379 AA;

QY 1 TPEKEPTAAPAEPEAPE 20  
:|||||:|:|:|:  
Db 221 SPESEEPSVAAPSEETPSE 240

RESULT 37  
ABU46857  
ID ABU46857 standard; protein; 379 AA.  
XX  
AC ABU46857;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #32384.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Streptococcus pyogenes.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US0009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA50727.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 74781; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required







This Page Blank (uspto)